



Global smoothness estimation of a Gaussian process from regular sequence designs

Delphine Blanke, Céline Vial

► To cite this version:

Delphine Blanke, Céline Vial. Global smoothness estimation of a Gaussian process from regular sequence designs. 2014. hal-00750409v2

HAL Id: hal-00750409

<https://hal.science/hal-00750409v2>

Preprint submitted on 8 Jan 2014

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

GLOBAL SMOOTHNESS ESTIMATION OF A GAUSSIAN PROCESS FROM REGULAR SEQUENCE DESIGNS

DELPHINE BLANKE AND CÉLINE VIAL

ABSTRACT. We consider a real Gaussian process X having a global unknown smoothness (r_0, β_0) , $r_0 \in \mathbb{N}_0$ and $\beta_0 \in]0, 1[$, with $X^{(r_0)}$ (the mean-square derivative of X if $r_0 \geq 1$) supposed to be locally stationary with index β_0 . From the behavior of quadratic variations built on divided differences of X , we derive an estimator of (r_0, β_0) based on - not necessarily equally spaced - observations of X . Various numerical studies of these estimators exhibit their properties for finite sample size and different types of processes, and are also completed by two examples of application to real data.

1. INTRODUCTION

In many areas assessing the regularity of a Gaussian process represents still and always an important issue. In a straightforward way, it allows to give accurate estimates for approximation or integration of sampled process. An important example is the kriging, which becomes more and more popular with growing number of applications. This method consists in interpolating a Gaussian random field observed only in few points. Estimating the covariance function is often the first step before plug this estimates in the Kriging equations, see Stein (1999). Usually the covariance function is assumed to belong to a parametric family, where these unknown parameters are linked to the sampled path regularity: for example the power model, which corresponds to a Fractional Brownian motion. Actually, many applications make use of irregular sampling and Stein (1999) (chap. 6.9) gives an hint of how adding three points very near from the origin among the already twenty equally spaced observations improve drastically the estimation of the regularity parameter. In this paper, we defined an estimator of global regularity of a Gaussian process when the sampling design is regular, that is observation points correspond to quantile of some distribution, see section 2 for details. Taking into account a non uniform design is innovating regarding other existing estimators and makes sense as to the remark above.

A wide range of methods have been proposed to reconstruct a sample path from discrete observations. For processes satisfying to the so-called Sacks and Ylvisaker (SY) conditions, recent works include: Müller-Gronbach (1996, orthogonal projection, optimal designs), Müller-Gronbach and Ritter (1997, linear interpolation, optimal designs), Müller-Gronbach and Ritter (1998, linear interpolation, adaptive designs). Under Hölder type conditions, one may cite e.g. works of Seleznev (1996, linear interpolation), Seleznev (2000, Hermite interpolation splines, optimal designs), Seleznev and Buslaev (1998, best approximation order). Note that a more detailed survey may be found in the book by Ritter (2000). Another important topic, involving the knowledge of regularity and arising in above cited works, is the search of an optimal design. In time series context, Cambanis (1985) analyzes three important problems (estimation of regression coefficients, estimation

Key words and phrases. Inference for Gaussian processes; Locally stationary process; Divided differences; Quadratic variations.

of random integrals and detection of signal in noise) for which he is looking for optimal design. The latter two problems involve approximations of integrals, where knowledge of process regularity is particularly important (see, e.g. Ritter, 2000), we provide a detailed discussion on this topic in section 4 together with additional references. Applications of estimation of regularity can be also find in Adler (1990) where bounds of suprema distributions depend on the sample roughness, in Istas (1992) where the regularity is involved in the choice of the best wavelet base in image analysis or more generally in prediction area, see Cuzick (1977); Lindgren (1979); Bucklew (1985).

Furthermore for a real stationary and non differentiable Gaussian process with covariance $K(s, t) = K(|t - s|, 0)$ such that $K(t, 0) = K(0, 0) - A|t|^{2\beta_0} + o(|t|^{2\beta_0})$ as $|t| \rightarrow 0$, the parameter β_0 , $0 < \beta_0 < 1$, is closely related to fractal dimension of the sample paths. This relationship is developed in particular in the works by Adler (1981) and Taylor and Taylor (1991) and it gave rise to an important literature around estimation of β_0 . Note that this relation can be extended, e.g. for non Gaussian process in Hall and Roy (1994). The recent paper of Gneiting et al. (2012) gives a review on estimator of the fractal dimension for times series and spatial data. They also provide a wide range of application in environmental science, e.g. hydrology, topography of sea floor. Note that this paper is restricted to the case of $(n + 1)$ equally spaced observations. In relation with our work, we refer especially to Constantine and Hall (1994) for estimators based on quadratic variations and their extensions developed by Kent and Wood (1997). Still for this stationary framework, Chan et al. (1995) introduce a periodogram-type estimator whereas Feuerverger et al. (1994) use the number of level crossings.

In this paper, our aim is to estimate the global smoothness (r_0, β_0) of a Gaussian process X , supposed to be r_0 -times differentiable (for some nonnegative integer r_0) where $X^{(r_0)}$ (the r_0 -th mean-square derivative of X for non-zero r_0) is supposed to be locally stationary with regularity β_0 . The parameters (r_0, β_0) being both unknown, we improve the previous works in several ways:

- not necessarily equally spaced observations of X over some finite interval $[0, T]$ are considered,
- X is not supposed to be stationary not even with stationary increments,
- X has an unknown degree of differentiability, r_0 , to be estimated,
- for $r_0 \geq 1$, the coefficient of smoothness β_0 is related to the unobserved derivative $X^{(r_0)}$.

Our methodology is based on an estimator of r_0 , say \hat{r}_0 , derived from quadratic variations of divided differences of X and consequently, generalize the estimator studied by Blanke and Vial (2011) for the equidistant case. In a second step, we proceed to the estimation of β_0 , with an estimator $\hat{\beta}_0$ which can be viewed as a simplification of that studied, in the case $r_0 = 0$, by Kent and Wood (1997). Also for processes with stationary increments and using a linear regression approach, Istas and Lang (1997) have proposed and studied an estimator of $H = 2(r_0 + \beta_0)$ for equally spaced observations. As far as we can judge, our two steps procedure seems to be simpler and more competitive. We obtain an upper bound for $\mathbb{P}(\hat{r}_0 \neq r_0)$ as well as the mean square error of \hat{r}_0 and almost sure rates of convergence of $\hat{\beta}_0$. Surprisingly, these almost sure rates are comparable to those obtained in the case of r_0 equal to 0: by this way the preliminary estimation of r_0 does not affect that of β_0 , even if $X^{(r_0)}$ is not observed. Next, in section 4, we derive theoretical and numerical results concerning the important application of approximation and integration. We complete this work with an extensive computational study: we compare different estimators of r_0 and $r_0 + \beta_0$ for processes with various kinds of smoothness, derive properties of our estimators for finite sample size, an example of consequence of the

misspecification of r_0 is given, and an example of process with trend is also study. To end this part, we apply our global estimation to two well-known real data sets: Roller data Laslett (1994) and Biscuit data Brown et al. (2001).

2. THE FRAMEWORK

2.1. The process and design. We consider a Gaussian process $X = \{X(t), t \in [0, T]\}$ observed at $(n + 1)$ instants on $[0, T]$, $T > 0$, with covariance function $\mathbb{K}(s, t) = \text{Cov}(X(s), X(t))$. We shall assume the following conditions on regularity of X .

Assumption 2.1 (A2.1). *X satisfies the following conditions.*

- (i) *There exists some nonnegative integer r_0 , such that X is r_0 -times differentiable in quadratic mean, denote $X^{(r_0)}$.*
- (ii) *The process $X^{(r_0)}$ is supposed to be locally stationary:*

$$\lim_{h \rightarrow 0} \sup_{s, t \in [0, T], |s-t| \leq h, s \neq t} \left| \frac{\mathbb{E}(X^{(r_0)}(s) - X^{(r_0)}(t))^2}{|s - t|^{2\beta_0}} - d_0(t) \right| = 0 \quad (2.1)$$

where $\beta_0 \in]0, 1[$ and d_0 is a positive continuous function on $[0, T]$.

- (iii-p) *For either $p = 1$ or $p = 2$, $\mathbb{K}^{(r_0+p, r_0+p)}(s, t)$ exists on $[0, T]^2 \setminus \{s = t\}$ and satisfies for some $D_p > 0$:*

$$\left| \mathbb{K}^{(r_0+p, r_0+p)}(s, t) \right| \leq D_p |s - t|^{-(2p-2\beta_0)}.$$

Moreover, we suppose that $\mu \in C^{r_0+1}([0, T])$.

Note that the local stationarity makes reference to Berman (1974)'s sense. The condition A2.1-(i) can be translated in terms of the covariance function. In particular the function \mathbb{K} is continuously differentiable with derivatives $\mathbb{K}^{(r, r)}(s, t) = \text{Cov}(X^{(r)}(s), X^{(r)}(t))$, for $r = 1, \dots, r_0$. Also, the mean of the process $\mu(t) := \mathbb{E}X(t)$ is a r_0 -times continuously differentiable function with $\mathbb{E}X^{(r)}(t) = \mu^{(r)}(t)$, $r = 0, \dots, r_0$. Conditions A2.1-(iii-p) are more technical but classical ones when estimating regularity parameter, see Constantine and Hall (1994); Kent and Wood (1997).

These assumptions are satisfied by a wide range of examples, e.g. the r_0 -fold integrated fractional Brownian motion or the Gaussian process with Matérn covariance function, i.e. $\mathbb{K}(t, 0) = \frac{\pi^{1/2} \phi}{2^{(\nu)-1} \Gamma(\nu+1/2)} (\alpha|t|)^\nu K_\nu(\alpha|t|)$, where K_ν is a modified Bessel function of the second kind of order ν . The latter process gets a global smoothness equal to $(\lfloor \nu \rfloor, \nu - \lfloor \nu \rfloor)$, see Stein (1999) p.31. Detailed examples, including different classes of stationary processes, can be found in Blanke and Vial (2008, 2011).

Note that, for processes with stationary increments, the function d_0 is reduced to a constant. Of course, cases with non constant $d_0(\cdot)$ are allowed as well as processes with a smooth enough trend. In particular, for some sufficiently smooth functions a and m on $[0, T]$, the process $Y(t) = a(t)X(t) + m(t)$ will also fulfills Assumption A2.1, see lemma 6.1 for details.

Let us turn now to the description of observation points. We consider that the process X is observed at $(n + 1)$ instants, denoted by

$$0 = t_{0,n} < t_{1,n} < \dots < t_{n,n} \leq T$$

where the $t_k := t_{k,n}$ form a regular sequence design. That is, they are defined as quantiles of a fixed positive and continuous density ψ on $[0, T]$:

$$\int_0^{t_k} \psi(s) ds = \frac{k \delta_n}{T}, \quad k = 0, \dots, n,$$

for δ_n a positive sequence such that $\delta_n \rightarrow 0$ and $n\delta_n \rightarrow T(-)$. Clearly, if ψ is the uniform density on $[0, T]$, one gets the equidistant case. Some further assumptions on ψ are needed to get some control over the t_k 's.

Assumption 2.2 (A2.2). *The density ψ satisfies:*

- (i) $\inf_{t \in [0, T]} \psi(t) > 0$,
- (ii) $\forall (s, t) \in [0, T]^2, |\psi(s) - \psi(t)| \leq L |s - t|^\alpha$, for some $\alpha \in]0, 1]$.

These hypothesis ensure a controlled spacing between two distinct points of observation, see Lemma 6.2. From a practical point of view, this flexibility may allow to recognize inhomogeneities in the process (e.g. presence of pics in environmental pollution monitoring, see Gilbert (1987) and references therein) or else to describe situations where data are collected at equidistant times but become irregularly spaced after some screening (see for example the wolfcamp-aquifer data in Cressie (1993) p. 212).

2.2. The methodology. In this part, we want to give background ideas about the construction of our estimate of the global regularity (r_0, β_0) , when the process is observed on a non equidistant grid. The main idea is to introduce divided differences, quantities generalizing the finite differences, studied by Blanke and Vial (2011, 2012). Let us first recall that the unique polynomial of degree r that interpolates a function g at $r + 1$ points t_k, \dots, t_{k+ur} (for some positive integer u) can be written as:

$$g[t_k] + g[t_k, t_{k+u}](t - t_k) + g[t_k, t_{k+u}, t_{k+2u}](t - t_k)(t - t_{k+u}) + \dots + g[t_k, \dots, t_{k+ru}](t - t_k) \dots (t - t_{k+(r-1)u}) \quad (2.2)$$

where the divided differences $g[\dots]$ are defined by $g[t_k] = g(t_k)$ and for $j = 1, \dots, r$ (using the Lagrange's representation):

$$g[t_k, \dots, t_{k+ju}] = \sum_{i=0}^j \frac{g(t_{k+iu})}{\prod_{m=0, m \neq i}^j (t_{k+iu} - t_{k+mu})}.$$

In particular, we write $g[t_k, \dots, t_{k+ru}] = \sum_{i=0}^r b_{ikr}^{(u)} g(t_{k+iu})$ with

$$b_{ikr}^{(u)} := \frac{1}{\prod_{m=0, m \neq i}^r (t_{k+iu} - t_{k+mu})}. \quad (2.3)$$

These coefficients are of particular interest. In fact their first non-zero moments are of order r . We can also derive an explicit bound and an asymptotic expansion for $b_{ikr}^{(u)}$, see lemma 6.3 for details.

Then, for positive integers r and u , we consider the u -dilated divided differences of order r for X :

$$D_{r,k}^{(u)} X = \sum_{i=0}^r b_{ikr}^{(u)} X(t_{k+iu}), \quad k = 0, \dots, n - ur \quad (2.4)$$

with $b_{ikr}^{(u)}$ defined by (2.3). Note that, if $\psi(t) = T^{-1} \mathbf{1}_{[0, T]}(t)$, the sequence of designs is equidistant, that is $t_{k,n} = k\delta_n$, and divided differences turn to be finite differences. More precisely, for the sequence $a_{i,r} = \binom{r}{i} (-1)^{r-i}$, let define the finite differences $\Delta_{r,k}^{(u)} = \sum_{i=0}^r a_{i,r} X((k + iu)\delta_n)$. Noticing that in the case of equally spaced observations,

$$b_{ikr}^{(u)} = \frac{(u\delta_n)^{-r}}{\prod_{m=0, m \neq i}^r (i - m)} = \frac{a_{i,r}}{r!} (u\delta_n)^{-r},$$

we deduce the relation $D_{r,k}^{(u)} X = \frac{(u\delta_n)^{-r}}{r!} \Delta_{r,k}^{(u)} X$. From now on, we set the u -diluted quadratic variations of X of order r : $\overline{(D_r^{(u)} X)^2} = \frac{\sum_{k=0}^{n_r} (D_{r,k}^{(u)} X)^2}{n_r + 1}$ with $n_r := n - ur$. Construction of our estimators is based on following asymptotic properties concerning the mean behavior of $\overline{(D_r^{(u)} X)^2}$:

$$(P) \begin{cases} \text{the quantity } \mathbb{E} \left(\overline{(D_r^{(u)} X)^2} \right) \text{ is of order } (u/n)^{-2(p-\beta_0)} \text{ for } r = r_0 + p \\ \text{for } p = 1, 2, \text{ and gets a finite non zero limit when } r \leq r_0. \end{cases}$$

See proposition 6.1 for precise results. These results imply that a good choice of r (namely $r = r_0 + 1$ or $r_0 + 2$) could provide an estimate of β_0 , at least with an adequate combination of u -diluted quadratic variations of X . To this end, we propose a two steps procedure:

Step 1: *Estimation of r_0 .*

Based on $D_{r,k}^{(1)} X$, we estimate r_0 with

$$\hat{r}_0 = \min \left\{ r \in \{2, \dots, m_n\} : \overline{(D_r^{(1)} X)^2} \geq n^2 b_n \right\} - 2. \quad (2.5)$$

If the above set is empty, we fix $\hat{r}_0 = l_0$ for an arbitrary value $l_0 \notin \mathbb{N}_0$. Here, $m_n \rightarrow \infty$ but if an upper bound B is known for r_0 , one has to choose $m_n = B + 2$. The threshold b_n is a positive sequence chosen such that : $n^{-2(1-\beta_0)} b_n \rightarrow 0$ and $n^{2\beta_0} b_n \rightarrow \infty$ for all $\beta_0 \in]0, 1[$. For example, omnibus choices are given by $b_n = (\ln n)^\alpha$, $\alpha \in \mathbb{R}$.

Step 2: *Estimation of β_0 .*

Next, we derive two families of estimators for β_0 , namely $\hat{\beta}_n^{(p)}$, with either $p = 1$ or $p = 2$ and u, v ($u < v$) given integers:

$$\hat{\beta}_n^{(p)} := \hat{\beta}_n^{(p)}(u, v) = p + \frac{1}{2} \frac{\ln \left(\overline{(D_{\hat{r}_0+p}^{(u)} X)^2} \right) - \ln \left(\overline{(D_{\hat{r}_0+p}^{(v)} X)^2} \right)}{\ln(u/v)}.$$

Remark 2.1 (The case of $r_0 = 0$ with equally spaced observations). *Kent and Wood (1997) proposed estimators of $2\beta_0 = \alpha$ based on ordinary and generalized least squares on the logarithm of the quadratic variations versus logarithm of a vector of values u , more precisely*

$$\hat{\alpha}^{(p)} = \frac{(\mathbf{1}^\top W \mathbf{1})(\mathbf{u}^\top W \mathbf{Q}^{(p)}) - (\mathbf{1}^\top W \mathbf{u})(\mathbf{1}^\top W \mathbf{Q}^{(p)})}{(\mathbf{1}^\top W \mathbf{1})(\mathbf{u}^\top W \mathbf{u}) - (\mathbf{1}^\top W \mathbf{u})^2}$$

where $\mathbf{1}$ is the m -vector of 1s, $\mathbf{u} = (\ln(u), u = 1, \dots, m)^\top$, $\mathbf{Q}^{(p)} = (\ln(\overline{(\Delta_p^{(u)} X)^2}), u = 1, \dots, m)^\top$ and W is either the identity matrix I_m of order $m \times m$ or a matrix depending on (n, β_0) which converges to the asymptotic covariance of $n^{1/2}(\overline{(\Delta_p^{(u)} X)^2} - \mathbb{E}(\overline{(\Delta_p^{(u)} X)^2}))$. The ordinary least square estimator – corresponding to $W = I_m$ – is denoted by $\hat{\alpha}_{OLS}^{(p)}$, where p is adapted to the regularity of the process (supposed to be known in their work). The choice $p = 1$, with the sequence $(-1, 1)$, leads to the estimator studied by Constantine and Hall (1994). Remark that, for $(u, v) = (1, 2)$, one gets $\hat{\beta}_n^{(1)} = \hat{\alpha}_{OLS}^{(0)}$ and $\hat{\beta}_n^{(2)} = \hat{\alpha}_{OLS}^{(1)}$ but, even in this equidistant case, new estimators may be derived with other choices of (u, v) such as $(u, v) = (1, 4)$ (which seems to perform well, see Section 5.2).

3. ASYMPTOTIC RESULTS

In Blanke and Vial (2011), an exponential bound is obtained for $\mathbb{P}(\hat{r}_0 \neq r_0)$ in the equidistant case, implying that, almost surely for n large enough, \hat{r}_0 is equal to r_0 . Here, we generalize this result to regular sequence designs but also, we complete it with the average behavior of \hat{r}_0 .

Theorem 3.1. *Under Assumption A2.1 (fulfilled with $p = 1$ or $p = 2$) and A2.2, we have $\mathbb{P}(\hat{r}_0 \neq r_0) = \mathcal{O}\left(\exp(-\varphi_n(p))\right)$ and $\mathbb{E}(\hat{r}_0 - r_0)^2 = \mathcal{O}\left(m_n^3 \exp(-\varphi_n(p))\right)$, where, for some positive constant $C_1(r_0)$, $\varphi_n(p)$ is defined by*

$$\varphi_n(p) = C_1(r_0) \times \begin{cases} n \mathbb{1}_{]0, \frac{1}{2}[}(\beta_0) + n(\ln n)^{-1} \mathbb{1}_{\{\frac{1}{2}\}}(\beta_0) + n^{2-2\beta_0} \mathbb{1}_{] \frac{1}{2}, 1[}(\beta_0) & \text{if } p = 1 \\ n & \text{if } p = 2. \end{cases}$$

Remark that one may choose m_n tending to infinity arbitrary slowly. Indeed, the unique restriction is that r_0 belongs to the grid $\{1, \dots, m_n\}$ for n large enough. From a practical point of view, one may choose a preliminary fixed bound B , and, in the case where the estimator return the non-integer value l_0 , replace B by B' greater than B .

The bias of $\hat{\beta}_n^{(p)}$ will be controlled by a second-order condition of local stationarity, more specifically we have to strengthen the relation (2.1) in:

$$\lim_{h \rightarrow 0} \sup_{\substack{s, t \in [0, T], \\ |s-t| \leq h, \\ s \neq t}} \left| |s-t|^{-\beta_1} \left(\frac{\mathbb{E}(X^{(r_0)}(s) - X^{(r_0)}(t))^2}{|s-t|^{2\beta_0}} - d_0(t) \right) - d_1(t) \right| = 0 \quad (3.1)$$

for a positive β_1 and continuous function d_1 .

Theorem 3.2. *If relation (3.1), Assumption A2.1 with $p = 1$ or $p = 2$, and A2.2 are fulfilled, we obtain*

$$\limsup_{n \rightarrow \infty} V_n^{(p)} \left| \hat{\beta}_n^{(p)} - \beta_0 \right| \leq C_1(p) \quad a.s.$$

where $C_1(p)$ is some positive constant and

$$V_n^{(1)} = \min \left(n^{\beta_1}, \sqrt{\frac{n}{\ln n}} \mathbb{1}_{]0, \frac{3}{4}[}(\beta_0) + \frac{\sqrt{n}}{\ln n} \mathbb{1}_{\{\frac{3}{4}\}}(\beta_0) + \frac{n^{2(1-\beta_0)}}{\ln n} \mathbb{1}_{] \frac{3}{4}, 1[}(\beta_0) \right),$$

$$V_n^{(2)} = \min \left(n^{\beta_1}, \sqrt{\frac{n}{\ln n}} \right).$$

Remark 3.1 (Rates of convergence with equally spaced observations). *For stationary gaussian processes, Kent and Wood (1997) give the mean square error and convergence in distribution of their estimator described in remark 2.1. They obtained the same rate up to a logarithmic order, due here to almost sure convergence, for both families $p = 1$ and $p = 2$. The asymptotic distribution is either of Gaussian or of Rosenblatt type depending on β_0 less or greater than $3/4$. Istas and Lang (1997) introduced an estimator of $H = 2(r_0 + \beta_0)$ (for stationary increment processes) with a global linear regression approach, based on an asymptotic equivalent of the quadratic variation and using some adequate family of sequences. For $r_0 = 0$, their approach matches with the previous one, with $p = 1$, in using dilated sequence of type $a_{jr} = \binom{r}{j}(-1)^{r-j}$. Assuming a known upper bound on r_0 , they derived convergence in distribution to a centered Gaussian variable with rate depending on β_0 -root- n for $\beta_0 \leq 3/4$ and $n^{1/2-\alpha(2\beta_0-3/2)}$, where $\delta_n = n^{-\alpha}$ and $\alpha < 1$, for $\beta_0 > 3/4$. In this last case, to obtain a gaussian limit they have to assume that r_0 is known, the observation interval is no more bounded, and the rate of convergence is lower than root- n .*

4. APPROXIMATION AND INTEGRATION

4.1. Results for plug-in estimation. A classical and interesting topic is approximation and/or integration of a sampled path. An extensive literature may be found on these topics with a detailed overview in the recent monograph of Ritter (2000). The general framework is as follows : let $X = \{X_t, t \in [0, 1]\}$, be observed at sampled times $t_{0,n}, \dots, t_{n,n}$ over $[a, b]$, more simply denoted by t_0, \dots, t_n . Approximation of $X(\cdot)$ consists in interpolation of the path on $[a, b]$, while weighted integration is the calculus of $\mathcal{I}_\rho = \int_a^b X(t)\rho(t)dt$ for some positive and continuous weight function ρ . These problems are closely linked, see e.g. Ritter (2000) p. 19-21. Closely to our framework of local stationary derivatives, we may refer more specifically to works of Plaskota et al. (2004) for approximation and Benhenni (1998) for integration. For sake of clarity, we give a brief summary of their obtained results. In the following, we denote by $\mathcal{H}(r_0, \beta_0)$ the family of Gaussian processes having r_0 derivatives in quadratic mean and r_0 -th derivative with Hölderian regularity of order $\beta_0 \in]0, 1[$. For measurable $g_i(\cdot)$, we consider the approximation $\mathcal{A}_{n,g}(t) = \sum_{i=0}^n X(t_i)g_i(t)$ and the corresponding weighted and integrated L^2 -error $e_\rho(\mathcal{A}_{n,g})$ with $e_\rho^2(\mathcal{A}_{n,g}) = \int_a^b \mathbb{E}|X(t) - \mathcal{A}_{n,g}(t)|^2 \rho(t)dt$. For $X \in \mathcal{H}(r_0, \beta_0)$ and known (r_0, β_0) , Plaskota et al. (2004) have shown that

$$0 < c(r_0, \beta_0) \leq \liminf_{n \rightarrow \infty} n^{r_0 + \beta_0} \inf_g e_\rho(\mathcal{A}_{n,g}) \leq \overline{\lim}_{n \rightarrow \infty} n^{r_0 + \beta_0} \inf_g e_\rho(\mathcal{A}_{n,g}) \leq C(r_0, \beta_0) < +\infty$$

for equidistant sampled times t_1, \dots, t_n and Gaussian processes defined and observed on the half-line $[0, +\infty[$. Of course, optimal choices of functions g_i , giving a minimal error, depend on the unknown covariance function of X .

For weighted integration, the quadrature is denoted by $\mathcal{Q}_{n,d} = \sum_{i=0}^n X(t_i)d_i$ with well-chosen constants d_i (typically, one may take $d_i = \int_a^b g_i(t)dt$). For known (r_0, β_0) , a short list of references could be:

- Sacks and Ylvisaker (1968, 1970) with $r_0 = 0$ or 1 , $\beta_0 = \frac{1}{2}$ and known covariance,
- Benhenni and Cambanis (1992) for arbitrary r_0 and $\beta_0 = \frac{1}{2}$,
- Stein (1995) for stationary processes and $r_0 + \beta_0 < \frac{1}{2}$,
- Ritter (1996) for minimal error, under Sacks and Ylvisaker's conditions, and with arbitrary r_0 .

Let us set $e_\rho^2(\mathcal{Q}_{n,d}) = \mathbb{E}|I_\rho - \mathcal{Q}_{n,d}|^2$, the mean square error of integration. In the stationary case and for known r_0 , Benhenni (1998) established the following exact behavior: If $\rho \in C^{r_0+3}([a, b])$ then for some given quadrature $\mathcal{Q}_{n,d^*(r_0)}$ on $[a, b]$,

$$n^{r_0 + \beta_0 + \frac{1}{2}} e_\rho(\mathcal{Q}_{n,d^*(r_0)}) \xrightarrow[n \rightarrow \infty]{} c_{r_0, \beta_0} \left(\int_a^b \rho^2(t) \psi^{-(2(r_0 + \beta_0) + 1)}(t) dt \right)^{\frac{1}{2}}$$

where ψ is the density relative to the regular sampling $\{t_1, \dots, t_n\}$. Moreover, following Ritter (1996), it appears that this last result is optimal under Sacks and Ylvisaker's conditions. Finally, Istas and Laredo (1997) have proposed a quadrature, requiring only an upper bound on r_0 , also with an error of order $\mathcal{O}(n^{-(r_0 + \beta_0 + \frac{1}{2})})$.

All these results shown the importance of well estimating r_0 and motivate ourself to focus on plugged-in interpolators, namely those using Lagrange polynomial of order estimated by \hat{r}_0 . More precisely, Lagrange interpolation of order $r \geq 1$ is

defined by

$$\tilde{X}_r(t) = \sum_{i=0}^r L_{i,k,r}(t) X(t_{kr+i}), \text{ with } L_{i,k,r}(t) = \prod_{\substack{j=0 \\ j \neq i}}^r \frac{(t - t_{kr+j})}{t_{kr+i} - t_{kr+j}}, \quad (4.1)$$

for $t \in \mathcal{I}_k := [t_{kr}, t_{kr+r}]$, $k = 0, \dots, \lfloor \frac{n}{r} \rfloor - 2$ and $\mathcal{I}_{\lfloor \frac{n}{r} \rfloor - 1} = [\lfloor \lfloor \frac{n}{r} \rfloor - 1 \rfloor r, T]$.

Our plugged method will consist in the approximation given by $\mathcal{A}_{n,L}(t) = \tilde{X}_{\max(\hat{r}_0, 1)}(t)$, $t \in [0, T]$, and quadrature by $\mathcal{Q}_{n,L} = \int_0^T \tilde{X}_{\hat{r}_0+1}(t) \rho(t) dt$. Indeed, Lagrange polynomials are of easy implementation and by the result of Plaskota et al. (2004) with known r_0 , they reach the optimal rate of approximation, $n^{-(r_0+\beta_0)}$ without requiring knowledge of covariance. Our following result shows also that the associate quadrature has the expected rate $n^{-(r_0+\beta_0+\frac{1}{2})}$. Indeed, in the weighted case and for $T > 0$, we obtain the following asymptotic bounds in the case of a regular design.

Theorem 4.1. *Suppose that conditions A2.1(i)-(ii) and A2.2 hold, choose a logarithmic order for m_n in (2.5) and consider a positive and continuous weight function ρ .*

(a) *Under condition A2.1(iii-1), we have*

$$e_\rho(\text{app}(\hat{r}_0)) := \left(\int_0^T \mathbb{E} \left| X(t) - \tilde{X}_{\max(\hat{r}_0, 1)}(t) \right|^2 \rho(t) dt \right)^{1/2} = \mathcal{O}(n^{-(r_0+\beta_0)}),$$

(b) *if condition A2.1(iii-2) holds:*

$$e_\rho(\text{int}(\hat{r}_0)) := \left(\mathbb{E} \left| \int_0^T (X(t) - \tilde{X}_{\hat{r}_0+1}(t)) \rho(t) dt \right|^2 \right)^{1/2} = \mathcal{O}(n^{-(r_0+\beta_0+\frac{1}{2})}).$$

In conclusion, expected rates for approximation and integration are reached by plugged Lagrange piecewise polynomials. Of course if r_0 is known, this last result holds true with \hat{r}_0 replaced by r_0 .

4.2. Simulation results. The figure 4.1 is obtained using 1000 simulated sample paths observed in equally spaced points on $[0, 1]$. This figure illustrates results of approximation for different processes. The logarithm of empirical integrated mean square error (in short IMSE), i.e. $e_1^2(\text{app}(\hat{r}_0))$, is drawn as a function of $\ln(n)$ with a range of sample size from 25 to 1000. We may notice that we obtain straight lines with slope very near to $-H = -2(r_0 + \beta_0)$. Since the Ornstein-Uhlenbeck process is a scaled time-transformed Wiener process, intercepts are different contrary to stationary versus non-stationary continuous ARMA processes.

5. NUMERICAL RESULTS

In this section, to numerically compare our estimators with existing ones, we restrict ourselves to the equidistant case with the choice $\psi(t) = \frac{1}{T} \mathbb{1}_{[0, T]}(t)$. As noticed before, we get for $a_{i,r} = \binom{r}{i} (-1)^{r-i}$ and $\Delta_{r,k}^{(u)} = \sum_{i=0}^r a_{i,r} X((k + iu)\delta_n)$, the relation $D_{r,k}^{(u)} X = \frac{(u\delta_n)^{-r}}{r!} \Delta_{r,k}^{(u)} X$ implying in turn that

$$\hat{H}_n^{(p)} = \frac{\ln \left(\left(\Delta_{\hat{r}_0+p}^{(u)} X \right)^2 \right) - \ln \left(\left(\Delta_{\hat{r}_0+p}^{(v)} X \right)^2 \right)}{\ln(u/v)} \quad (5.1)$$

is a consistent estimator of $H = 2(r_0 + \beta_0)$. All the simulation results are obtained by simulation of trajectories using two different methods : for stationary processes or with stationary increments we use the procedure described in Wood and Chan (1994) and for CARMA (continuous ARMA) processes, we use Tsai and Chan

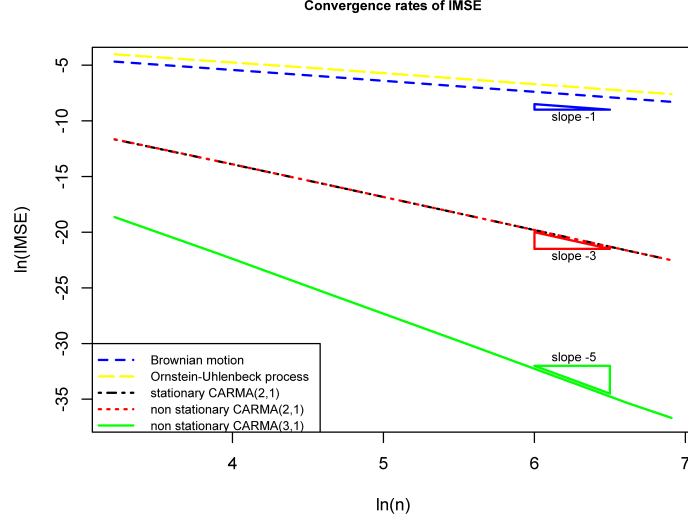


FIGURE 4.1. Logarithm of $e_1^2(\text{app}(\hat{r}_0))$, i.e. the IMSE in function of $\ln(n)$, for different processes. The dashed line corresponds to Brownian motion, long dashed line to O.U., dashed line to non stationary CARMA(2,1), dotted-dashed line to stationary CARMA(2,1) and solid line to non stationary CARMA(3,1). The small triangles near lines are here to indicate the theoretic slope.

(2000). Each of them consists in n equally spaced observation points on $[0, 1]$ and 1000 simulated sample paths. All computations have been performed with the R software (R Core Team, 2012).

5.1. Results for estimators of r_0 . This section is dedicated to the numerical properties of two estimators of r_0 . We consider the estimator introduced by Blanke and Vial (2011), derived from (2.5) in the equidistant case. An alternative, says \tilde{r}_n , based on Lagrange interpolator polynomials was proposed by Blanke and Vial (2008). More precisely, for $\delta_n = n^{-1}$ et $T = 1$, \tilde{r}_n is defined by

$$\tilde{r}_n = \min \left\{ r \in \{1, \dots, m_n\} : \frac{1}{r \tilde{n}_r} \sum_{k=0}^{r \tilde{n}_r - 1} \left(X\left(\frac{2k+1}{n}\right) - \tilde{X}_r\left(\frac{2k+1}{n}\right) \right)^2 \geq n^{-2r} b_n \right\} - 1$$

where $\tilde{n}_r = \lfloor \frac{n}{2r} \rfloor$ and $\tilde{X}_r(s)$ is defined for all $s \in [0, 1]$ and each $r \in \{1, \dots, m_n\}$ in the following way : there exist $k = 0, \dots, \tilde{n}_r - 1$ such that for $t \in \mathcal{I}_{2k} := [\frac{2kr}{n}, \frac{2(k+1)r}{n}]$, the piecewise Lagrange interpolation of $X(t)$, $\tilde{X}_r(t)$, is given by $\tilde{X}_r(t) = \sum_{i=0}^r L_{i,k,r}(t) X((kr+i)n^{-1})$, with $L_{i,k,r}(t) = \prod_{j=0, j \neq i}^r \frac{(t - (kr+j)n^{-1})}{(i-j)n^{-1}}$.

Both estimators use the critical value b_n which is involved in detection of the jump. Here, due to convergence properties, we make the choice $b_n = (\ln n)^{-1}$. Table 5.1 illustrates the strong convergence of both estimators and shows that this convergence is valid even for small number of observation points n , up to 10 for the estimator \hat{r}_0 . We may noticed that, in the case of bad estimation, our estimators overestimate the number of derivatives. Remark also that, for identical sample paths, \hat{r}_0 seems to be more robust than \tilde{r}_n . This behavior was expected as the latter uses only half of the observations for the detection of the jump in quadratic

TABLE 5.1. Value of the empirical probability that \hat{r}_0 or \tilde{r}_n equals r_0 or $r_0 + 1$ with $n = 10$ or 25.

	Wiener process, $r_0 = 0$		CARMA(2,1), $r_0 = 1$		CARMA(3,1), $r_0 = 2$	
	Number of equally spaced observations n					
event	10	25	10	25	10	25
$\hat{r}_n = r_0$	0.995	1.000	0.913	1.000	0.585	0.999
$\tilde{r}_n = r_0 + 1$	0.005	0.000	0.087	0.000	0.415	0.001
$\hat{r}_0 = r_0$	1.000	1.000	1.000	1.000	0.999	1.000
$\hat{r}_0 = r_0 + 1$	0.000	0.000	0.000	0.000	0.001	0.000

TABLE 5.2. Value of the empirical probability that \hat{r}_0 or \tilde{r}_n equals r_0 for a fractional Brownian motion or an integrated one with fractal index $2\beta_0$.

	$\hat{r}_0 = r_0$					$\tilde{r}_n = r_0$				
	number of equally spaced observations n									
	50	100	500	1000	1200	50	100	500	1000	1200
fBm β_0										
0.90	1.000	1.000	1.000	1.000	1.000	0.655	0.970	1.000	1.000	1.000
0.95	0.969	0.999	1.000	1.000	1.000	0.002	0.002	0.004	0.134	0.331
0.97	0.242	0.521	1.000	1.000	1.000	0.000	0.000	0.000	0.000	0.000
0.98	0.019	0.015	0.0420	0.5258	0.759	0.000	0.000	0.000	0.000	0.000
ifBm β_0										
0.02	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.90	1.000	1.000	1.000	1.000	1.000	0.000	0.000	0.645	0.999	1.000
0.95	0.305	0.888	1.000	1.000	1.000	0.000	0.000	0.000	0.000	0.000
0.97	0.000	0.000	0.292	0.993	1.000	0.000	0.000	0.000	0.000	0.000

mean. In these first results, processes have fractal index β_0 equals to $1/2$, but alternative choices of β_0 are of interest, so we consider the fractional Brownian motion (in short fBm) and the integrated fractional Brownian motion (in short ifBm), with respectively $r_0 = 0$ and $r_0 = 1$ and various values of β_0 .

Table 5.2 shows that \hat{r}_0 succeeds in estimating the true regularity for β_0 up to 0.9. Of course the number of observations must be large enough and, even more important for large values of r_0 when $\beta_0 \geq 0.95$. This latter result is clearly apparent when one compares the errors obtained for an ifBm with $\beta_0 = 0.95$ and a fBm with $\beta_0 = 0.95$. Finally, we can see once more that \tilde{r}_n is less robust against increasing β_0 , whereas our simulations have shown that, for $n = 2000$ and each simulated path, the estimator \hat{r}_0 is able to distinguish processes with regularity $(0, 0.98)$ and $(1, 0.02)$, an almost imperceptible difference!

5.2. Estimation of H and β_0 . This part is dedicated to the numerical properties of estimators $\hat{H}_n^{(p)}$, for $p = 1$ or 2 using the values $u = 1$ and $v = 4$ (giving more homogeneous results than $u = 1$ and $v = 2$). It ends with real data examples.

5.2.1. Quality of estimation. For the numerical part, we focus on the study of fBm, ifBm and, CARMA(3,1) with $r_0 = 2$, $\beta_0 = 0.5$. Table 5.3 illustrates the performance of our estimators when β_0 , r_0 are increasing: we compute the empirical mean-square error from our 1000 simulated sample paths and $n = 1000$ equally spaced observations are considered. It appears that, contrary to $\hat{H}_n^{(2)}$, the estimator $\hat{H}_n^{(1)}$ slightly deteriorates for values of β_0 greater than 0.8. This result is in agreement with the rate of convergence of Theorem 3.2, that depends on β_0 for this estimator.

TABLE 5.3. Values of mean square error and bias (between brackets) for estimators $\hat{H}_n^{(p)}$, for $p = 1$ or 2 and $n = 1000$.

	fBm β_0				
	0.2	0.5	0.8	0.9	0.95
$\hat{H}_n^{(1)}$	0.0017 (-0.0020)	0.0019 (-0.0045)	0.0034 (-0.0120)	0.0054 (-0.0295)	0.0065 (-0.0448)
$\hat{H}_n^{(2)}$	0.0030 (-0.0026)	0.0039 (-0.0038)	0.0040 (-0.0057)	0.0039 (-0.0069)	0.0039 (-0.0084)
	ifBm β_0				
	0.2	0.5	0.8	0.9	0.95
$\hat{H}_n^{(1)}$	0.0032 (-0.0072)	0.0026 (-0.0047)	0.0041 (-0.0150)	0.0061 (-0.0342)	0.0073 (-0.0488)
$\hat{H}_n^{(2)}$	0.0055 (-0.0106)	0.0051 (-0.0060)	0.0046 (-0.0060)	0.0044 (-0.0061)	0.0043 (-0.0061)

TABLE 5.4. Rates of convergence illustrated by linear regression for n in $\{500, 750, 1000, 1250\}$.

	$\hat{H}_n^{(1)}$		$\hat{H}_n^{(2)}$	
	slope	R^2	slope	R^2
fBm $\beta_0 = 0.5$	-0.488	0.998	-0.489	0.995
$\beta_0 = 0.6$	-0.475	0.998	-0.488	0.995
$\beta_0 = 0.7$	-0.426	0.994	-0.489	0.997
$\beta_0 = 0.8$	-0.334	0.989	-0.491	0.997
$\beta_0 = 0.9$	-0.225	0.990	-0.495	0.997
$\beta_0 = 0.95$	-0.186	0.995	-0.503	0.997
ifBm $\beta_0 = 0.9$	-0.302	0.987	-0.561	0.999
$\beta_0 = 0.95$	-0.244	0.978	-0.559	0.999

The bias is negative and seems to be unsensitive to the value of r_0 but the mean-square error is slightly deteriorated from $r_0 = 0$ to $r_0 = 1$ in both cases. Finally, for $\beta_0 < 0.8$, $\hat{H}_n^{(1)}$ seems preferable to $\hat{H}_n^{(2)}$, possibly due to a lower variance of this estimator. Nevertheless, both estimators perform globally well on these numerical experiments.

5.2.2. Asymptotic properties. Results of Theorem 3.2 are also illustrated in Table 5.4 where we have computed the regression of $\ln(\mathbb{E}|\hat{H}_n^{(p)} - H|)$ on $\ln n$ for various values of n and $\mathbb{E}|\hat{H}_n^{(p)} - H|$ estimated from our 1000 simulated sample paths. As expected, the slope (corresponding to our arithmetical rate of convergence) is constant and approximatively equal to 0.5 for $\hat{H}_n^{(2)}$ while, for $\hat{H}_n^{(1)}$, the decrease is apparent for high values of β_0 . Finally, Figure 5.1 illustrates the behavior of the estimators $\hat{H}_n^{(p)}$ with $p = 1$ or 2 , for different values of the regularity parameter β_0 . As we can see, boxplots deteriorates only slightly for $n = 100$ and 250 when β_0 increases from 0.5 to 0.8 but the dispersion for $\hat{H}_n^{(2)}$ is quite larger. For $\beta_0 = 0.95$, $\hat{H}_n^{(2)}$ clearly outperforms $\hat{H}_n^{(1)}$ with $n = 500$ observations. Estimation appears more difficult for smaller values of n , but it is a quite typical behavior in our considered framework.

5.2.3. Impact of misspecification of regularity. Next, Table 5.5 illustrates the impact of estimating H when the order r in quadratic variation is misspecified. In fact estimating β_0 requires the knowledge of r_0 or an upper bound of it. On the other hand, working with a too high value of r_0 may induce artificial variability in estimation, so a precise estimation of r_0 is important. Here, our numerical results

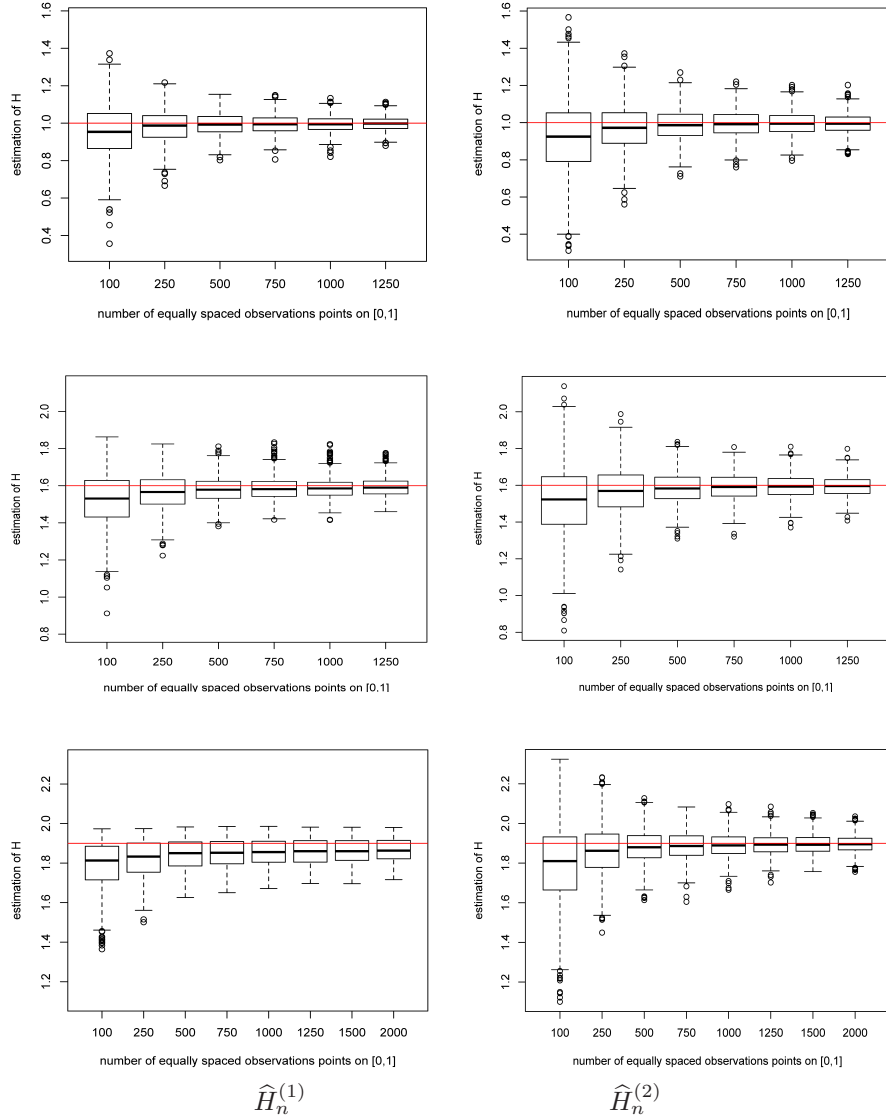


FIGURE 5.1. Each boxplot corresponds to 1000 estimations of H by $\hat{H}_n^{(1)}$ on the left and $\hat{H}_n^{(2)}$ on the right of the graph. Each realization consists in n equally spaced observations on $[0, 1]$ of a fBm with $\beta_0 = 0.5$ (top), $\beta_0 = 0.8$ (middle), $\beta_0 = 0.95$ (bottom), where $n = 100, 250, 500, 750, 1000, 1250, 1500, 2000$. The solid line corresponds to the real value of H .

show that, if the order r of quadratic variation used for estimating β_0 is less than $r_0 + 1$, then the quantity estimated is $2r$ and not $2(r + \beta_0)$.

5.2.4. Processes with varying trend or non constant function d_0 . All previous examples are locally stationary with a constant function d_0 . Processes meeting our conditions but with no stationary increments may be constructed using Lemma 6.1. As an example, from Y a standard Wiener process ($r_0 = 0, \beta_0 = 0.5$) or an integrated one ($r_0 = 1, \beta_0 = 0.5$), we simulate $X(t) = (t^{r_0+0.7} + 1)Y(t)$ having the regularity $(r_0, 0.5)$ and $d_0(t)$ equaling to $(t^{r_0+0.7} + 1)^2$. Figure 5.2 illustrates a Wiener sample

TABLE 5.5. Mean value and standard deviation (between brackets) of the estimator $\hat{H}_n^{(p)}$ based on quadratic variation of order r_0 or $r_0 - 1$ instead of $\hat{r}_0 + 1$ or $\hat{r}_0 + 2$.

	Order $r_0 - 1$		Order r_0	
	number n of equidistant observations			
	100	500	100	500
ifBm $\beta_0 = 0.2$			1.903 (0.064)	1.961 (0.025)
$\beta_0 = 0.5$			1.955 (0.035)	1.992 (0.006)
$\beta_0 = 0.8$			1.966 (0.024)	1.994 (0.004)
CARMA(3,1) ($r_0=2, \beta_0=0.5$)	1.970 (0.0140)	1.994 (0.003)	3.919 (0.058)	3.985 (0.0109)

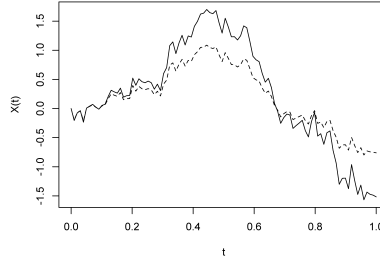


FIGURE 5.2. Wiener process (solid) and its locally stationary transformation (dashed) used in Table 5.6.

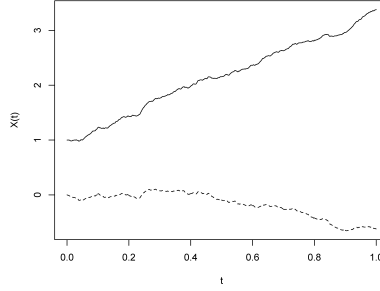


FIGURE 5.3. Sample path of a fBm with $\beta_0 = 0.8$ (dashed line) and the same with a trend $m(t) = (1+t)^2$ (solid line).

path and its transformation. Results are summarized in Table 5.6: comparing with Table 5.3 ($\beta_0 = 0.5$), it appears that the estimation is only slightly damaged for $r_0 = 1$ but of the same order when $r_0 = 0$. Other non stationary processes may also be obtained by adding some smooth trend. To this aim, we used same sample paths as for Table 5.3 with the additional trend $m(t) = (1+t)^2$, see Figure 5.3. We may noticed in Table 5.7 that we obtain exactly the same results for the estimator $\hat{H}_n^{(2)}$ and that only a slight loss is observed for $\hat{H}_n^{(1)}$.

5.3. Real data. Let us turn to examples based on real data sets. In this part, we compare our estimators of H with those proposed by Constantine and Hall (1994); Kent and Wood (1997). We compute estimated values by setting $(u, v) = (1, m)$

TABLE 5.6. Value of MSE and bias (between brackets) for non constant $d_0(\cdot)$.

	Wiener	Integrated Wiener
$\widehat{H}_n^{(1)}$	0.0021 (-0.0032)	0.0032 (0.0061)
$\widehat{H}_n^{(2)}$	0.0043 (-0.0042)	0.0058 (-0.0091)

TABLE 5.7. Value of MSE and bias (between brackets) for estimators $\widehat{H}_n^{(p)}$, for $p = 1$ or 2 in presence of a smooth trend.

	fBm		ifBm	
	$\beta_0 = 0.5$	$\beta_0 = 0.8$	$\beta_0 = 0.5$	$\beta_0 = 0.8$
$\widehat{H}_n^{(1)}$	0.0022 (0.0151)	0.0283 (0.1525)	0.0027 (0.0083)	0.0141 (0.0863)
$\widehat{H}_n^{(2)}$	0.0039 (-0.0038)	0.0040 (-0.0057)	0.0051 (-0.0060)	0.0046 (-0.0060)

in (5.1) with m in $\{2, 4, 6, 8, 10\}$ while for $\widehat{\alpha}_{OLS}^{(p)}$, $p = 1, 2$, defined in Remark 2.1 regression is carried out over $\mathbf{u} = (\ln(u), u = 1, \dots, m)^\top$.

5.3.1. *Roller data.* We first focus on roller height data introduced by Laslett (1994), which consists in $n = 1150$ heights measure at 1 micron intervals along a drum of a roller. This example was already studied in Kent and Wood (1997): they noticed that local self similarity may hold at sufficiently fine scales, so the regularity r_0 was supposed to be zero. Indeed, our estimator \widehat{r}_0 , directly used on the data with $b_n = 1/\ln(n)$, gives $\widehat{r}_0 = 0$ (with a value of $n^{4-2}(\Delta_2^{(1)}X)^2$ equal to 1172345). Next, we compute the values obtained for the estimation of H in Table 5.8, where values of estimates proposed by Constantine and Hall (1994); Kent and Wood (1997) are also reported for comparison. It should be observed that our simplified estimators present a similar sensitivity to the choice of m .

TABLE 5.8. Estimates in the roller height example

m	$\widehat{H}_n^{(1)}$	$\widehat{\alpha}_{OLS}^{(0)}$	$\widehat{H}_n^{(2)}$	$\widehat{\alpha}_{OLS}^{(1)}$
2	0.63	0.63	0.77	0.77
4	0.50	0.51	0.63	0.65
6	0.38	0.39	0.49	0.51
8	0.35	0.33	0.44	0.42
10	0.30	0.28	0.39	0.35

TABLE 5.9. Means of estimates in the biscuit example

	$m = 2$	$m = 4$	$m = 6$	$m = 8$	$m = 10$
$\widehat{H}_n^{(1)}$	3.60 (0.12)	3.67 (0.07)	3.65 (0.05)	3.62 (0.04)	3.59 (0.04)
$\alpha_{OLS}^{(1)}$	3.60 (0.12)	3.67 (0.07)	3.66 (0.05)	3.63 (0.04)	3.60 (0.03)
$\widehat{H}_n^{(2)}$	2.84 (0.45)	3.69 (0.30)	3.83 (0.24)	3.84 (0.19)	3.83 (0.16)
$\alpha_{OLS}^{(2)}$	2.84 (0.45)	3.67 (0.31)	3.91 (0.23)	3.98 (0.18)	3.99 (0.14)

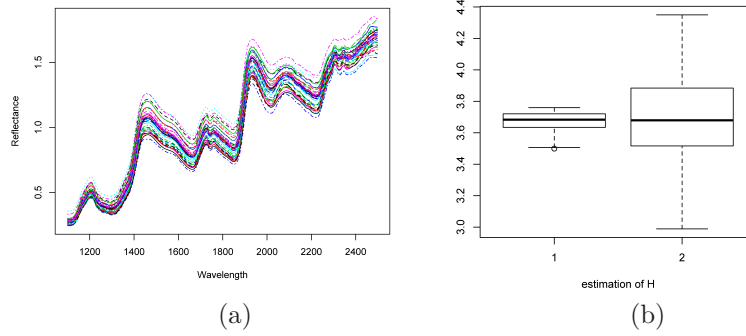


FIGURE 5.4. (a) Curve drawing reflectance in function of wavelength, varying between 1100 and 2498. (b) Box-plots for both estimators on the left $\hat{H}_n^{(1)}$, on the right $\hat{H}_n^{(2)}$ for the 39 curves and $(u, v) = (1, 4)$.

5.3.2. Biscuit data. Now, in order to compare the (empirical) variances of these estimators, we consider a second example introduced by Brown et al. (2001). The experiment involved varying the composition of biscuit dough pieces and data consist in near infrared reflectance (NIR) spectra for the same dough. The 40 curves are graphed on the figure 5.4. Each represents the near-infrared spectrum reflectance measure at each 2 nanometers from 1100 to 2498 nm, then 700 observation points for each biscuit. According to Brown et al. (2001), the observation 23 appears as an outlier. We estimate r_0 for each of the left 39 curves, using the threshold $b_n = 1$, which gives $\hat{r}_0 = 1$ for each curve. Furthermore, the averaged mean quadratic variation $n^{2r-2} (D_r^{(1)} X)^2$ equals to 0.33 when $r = 2$ and 122133 when $r = 3$, this explosion confirming the choice $\hat{r}_0 = 3 - 2 = 1$. We turn to estimation of H , having in mind the comparison of our estimators together with $\alpha_{OLS}^{(p)}$ (where $p = 1$ corresponds to the choice $(1, -2, 1)$ for a_{jr} and $p = 2$ to the choice $(-1, 3, -3, 1)$). The results are summarized in Table 5.9 where it appears that, for order $\hat{r}_0 + 2 = 3$, our estimator $\hat{H}_n^{(2)}$ seems to be less sensitive toward high values of m . Also our simplified estimators present a similar variance to $\hat{\alpha}_{OLS}^{(p)}$, $p = 1, 2$. To conclude this part, it should be noticed that for the 23rd curve, the choice $m = 4$ gives $\hat{H}_n^{(1)} = 3.64$ and $\hat{H}_n^{(2)} = 3.55$. It appears that, in both cases, these values belong to the interquartile range obtained from the 39 curves, so at least concerning the regularity, the curve 23 should not be considered as an outlier.

6. ANNEXES

6.1. Proofs of section 2.

Lemma 6.1. *Let Y be a zero mean process with given regularity (r_0, β_0) and asymptotic function $d_0(t) \equiv C_{r_0, \beta_0}$ that satisfies A2.1(iii-p) ($p = 1$ or 2). For a positive function $a \in C^{r_0+p}([0, T])$ and $m \in C^{r_0+p}([0, T])$, if $X(t) = a(t)Y(t) + m(t)$, then X has regularity (r_0, β_0) with asymptotical function $D_{r_0, \beta_0}(t) = a^2(t)C_{r_0, \beta_0}$ and satisfies A2.1(iii-p).*

Proof. See Seleznev (2000) and straightforward computation. \square

Lemma 6.2. *Under Assumptions A2.2, we get for $k = 0, \dots, n-1$, and $i = 1, \dots, n-k$:*

$$t_{k+i} - t_k \geq C_1 i \delta_n, \quad C_1 = (T \sup_{t \in [0, T]} \psi(t))^{-1}, \quad (6.1)$$

$$t_{k+i} - t_k \leq C_2 i \delta_n, \quad C_2 = (T \inf_{t \in [0, T]} \psi(t))^{-1}, \quad (6.2)$$

and if $i = 1, \dots, i_{\max}$ with i_{\max} not depending on n :

$$t_{k+i} - t_k = \frac{i \delta_n}{T \psi(t_k)} (1 + \mathcal{O}(\delta_n^\alpha)) \quad (6.3)$$

where $\mathcal{O}(\dots)$ is uniform over i and k .

Proof. Relations (6.1)-(6.2) are obtained with the mean-value theorem that induces, for $k = 0, \dots, n-1$ and $i = 1, \dots, n-k$:

$$t_{k+i} - t_k = \frac{i \delta_n}{T \psi(t_k + \theta(t_{k+i} - t_k))}, \quad 0 < \theta < 1.$$

To obtain the equivalence (6.3), one may write $t_{k+i} - t_k = \frac{i \delta_n}{T \psi(t_k)} (1 + R_n)$ with R_n defined by

$$R_n = \frac{\psi(t_k) - \psi(t_k + \theta(t_{k+i} - t_k))}{\psi(t_k + \theta(t_{k+i} - t_k))} \leq \frac{L |t_{k+i} - t_k|^\alpha}{\inf_{t \in [0, T]} \psi(t)} = \mathcal{O}(\delta_n^\alpha)$$

by Assumption A2.2 and uniformly over i , n and k for $i = 1, \dots, i_{\max}$. \square

Lemma 6.3. *We have under Assumption A2.2 and for $r \geq 1$, $i = 0, \dots, r$,*

(1) *for $p = 0, \dots, r-1$ and convention $0^0 = 1$*

$$\sum_{i=0}^r (t_{k+iu} - t_k)^p b_{ikr}^{(u)} = 0, \quad (6.4)$$

(2)

$$\sum_{i=0}^r (t_{k+iu} - t_k)^r b_{ikr}^{(u)} = 1, \quad (6.5)$$

(3)

$$|b_{ikr}^{(u)}| \leq \frac{C_1^{-r} u^{-r} \delta_n^{-r}}{\prod_{m=0, m \neq i}^r |i - m|}, \quad (6.6)$$

with C_1 given by (6.1),

(4)

$$b_{ikr}^{(u)} = \frac{u^{-r} \psi^r(t_k) T^r \delta_n^{-r}}{\prod_{m=0, m \neq i}^r (i - m)} (1 + \mathcal{O}(\delta_n^\alpha)) \quad (6.7)$$

with $\mathcal{O}(\dots)$ uniform over i and k .

Proof. The term $g[t_k, \dots, t_{k+ru}] = \sum_{i=0}^r b_{ikr}^{(u)} g(t_{k+iu})$ is the leading coefficient in the polynomial approximation of degree r of g , given in the decomposition (2.2). Considering the polynomial $g(t) = (t - t_k)^p$, we may immediately deduce the properties (6.4)-(6.5), from uniqueness of relation (2.2). Next, (6.6)-(6.7) are direct consequences of Lemma 6.2 and definition (2.3) of $b_{ikr}^{(u)}$. \square

Proposition 6.1. *Under Assumption A2.1 and A2.2, one obtains:*

(i) *for $r = r_0 + p$ with $p = 1, 2$:*

$$n^{-2(p-\beta_0)} \mathbb{E} \left(\overline{(D_{r_0+p}^{(u)} X)^2} \right) \xrightarrow{n \rightarrow \infty} u^{-2(p-\beta_0)} \ell(p, r_0, \beta_0)$$

where $\ell(p, 0, \beta_0) = -\frac{1}{2} \int_0^T d_0(t) \frac{\psi^{2p+1}(t)}{\psi^{2\beta_0}(t)} dt \sum_{i,j=0}^p \frac{|i-j|^{2\beta_0}}{\prod_{m=0, m \neq i}^p (i-m) \prod_{q=0, q \neq j}^p (j-q)}$ while if $r_0 \geq 1$,

$$\ell(p, r_0, \beta_0) = \frac{(-1)^{r_0+1} \int_0^T d_0(t) \frac{\psi^{2p+1}(t)}{\psi^{2\beta_0}(t)} dt}{2(2\beta_0 + 2r_0) \cdots (2\beta_0 + 1)} \sum_{i,j=0}^{r_0+p} \frac{|i-j|^{2(r_0+\beta_0)}}{\prod_{m=0, m \neq i}^{r_0+p} (i-m) \prod_{q=0, q \neq j}^{r_0+p} (j-q)} \quad (6.8)$$

(ii) for $r_0 \geq 1$ and $r = 1, \dots, r_0$:

$$\mathbb{E}\left(\overline{(D_r^{(u)} X)^2}\right) \xrightarrow{n \rightarrow \infty} \frac{1}{(r!)^2} \int_0^T \mathbb{E}(X^{(r)}(t))^2 \psi(t) dt.$$

Proof. A. Let us begin with general expressions of $\mathbb{E}\left(D_{r,k}^{(u)} X D_{r,\ell}^{(u)} X\right)$ useful for the sequel. First for $\mathbb{L}^{(p,p)}(s, t) = \mathbb{E}(X^{(p)}(s)X^{(p)}(t))$ ($p \geq 0$), the relation (2.1) is equivalent to

$$\lim_{h \rightarrow 0} \sup_{\substack{s, t \in [0, T] \\ |s-t| \leq h, s \neq t}} \left| \frac{\mathbb{L}^{(r_0, r_0)}(s, s) + \mathbb{L}^{(r_0, r_0)}(t, t) - 2\mathbb{L}^{(r_0, r_0)}(s, t)}{|s-t|^{2\beta_0}} - d_0(t) \right| = 0. \quad (6.9)$$

For $(v, w) \in [0, 1]^2$, we set $\dot{v}_{ik} = t_k + (t_{k+iu} - t_k)v$ and $\dot{w}_{j\ell} = t_\ell + (t_{\ell+ju} - t_\ell)w$. Next, from the definition of $D_{r,k}^{(u)} X$ given in (2.4), we get

$$\mathbb{E}(D_{r,k}^{(u)} X D_{r,\ell}^{(u)} X) = \sum_{i,j=0}^r b_{ikr}^{(u)} b_{j\ell r}^{(u)} \mathbb{L}^{(0,0)}(t_{k+iu}, t_{\ell+ju}).$$

For $r_0 = 0$ and since $\sum_{i=0}^r b_{ikr}^{(u)} = 0$, we have:

$$\begin{aligned} \mathbb{E}(D_{r,k}^{(u)} X D_{r,\ell}^{(u)} X) &= \sum_{i,j=0}^r b_{ikr}^{(u)} b_{j\ell r}^{(u)} \left\{ \mathbb{L}^{(0,0)}(t_{k+iu}, t_{\ell+ju}) \right. \\ &\quad \left. - \frac{1}{2} \mathbb{L}^{(0,0)}(t_{k+iu}, t_{k+iu}) - \frac{1}{2} \mathbb{L}^{(0,0)}(t_{\ell+ju}, t_{\ell+ju}) \right\}. \end{aligned} \quad (6.10)$$

If $r_0 \geq 1$, we apply multiple Taylor series expansions with integral remainder. Next, the properties $\sum_{i=0}^r b_{ikr}^{(u)} (t_{k+iu} - t_k)^p = 0$ for $p = 0, \dots, r-1$ (and convention $0^0 = 1$) induce :

$$\begin{aligned} \mathbb{E}(D_{r,k}^{(u)} X D_{r,\ell}^{(u)} X) &= \sum_{i,j=0}^r b_{ikr}^{(u)} b_{j\ell r}^{(u)} (t_{k+iu} - t_k)^{r^*} (t_{\ell+ju} - t_\ell)^{r^*} \\ &\quad \times \iint_{[0,1]^2} \frac{(1-v)^{r^*-1} (1-w)^{r^*-1}}{((r^*-1)!)^2} \mathbb{L}^{(r^*, r^*)}(\dot{v}_{ik}, \dot{w}_{j\ell}) dv dw \end{aligned} \quad (6.11)$$

where we have set $r^* = \min(r_0, r) \geq 1$.

B. From expressions (6.10)-(6.11), we are in position to derive the asymptotic behavior of $\mathbb{E}(\overline{(D_r^{(u)} X)^2})$.

Case $r_0 \geq 1$, $r = r_0 + p$, $p = 1$ or $p = 2$. In this case, $r^* = r_0 \leq r - 1$. From (6.11) and the property $\sum_{i=0}^r b_{ikr}^{(u)}(t_{k+iu} - t_k)^{r_0} = 0$, we may write

$$\begin{aligned} \mathbb{E}(D_{r,k}^{(u)} X)^2 &= \sum_{i,j=0}^r b_{ikr}^{(u)} b_{jkr}^{(u)} (t_{k+iu} - t_k)^{r_0} (t_{k+ju} - t_k)^{r_0} \int_0^1 \int_0^1 \frac{((1-v)(1-w))^{r_0-1}}{((r_0-1)!)^2} \\ &\quad \times \left\{ \mathbb{L}^{(r_0, r_0)}(\dot{v}_{ik}, \dot{w}_{jk}) - \frac{1}{2} \mathbb{L}^{(r_0, r_0)}(\dot{v}_{ik}, \dot{v}_{ik}) - \frac{1}{2} \mathbb{L}^{(r_0, r_0)}(\dot{w}_{jk}, \dot{w}_{jk}) \right\} dv dw \end{aligned} \quad (6.12)$$

Using the locally stationary condition (6.9), uniform continuity of $d_0(\cdot)$ on $[0, T]$ and the bound: $|\dot{v}_{ik} - \dot{w}_{jk}| \leq t_{k+ru} - t_k \leq C_1 r u \delta_n$ for $i = 0, \dots, r$ and $j = 0, \dots, r$, we may show that the predominant term for $\mathbb{E}((D_r^{(u)} X)^2)$ is given by:

$$\begin{aligned} \frac{-1}{2(n_r + 1)} \sum_{k=0}^{n_r} \sum_{i,j=0}^r b_{ikr}^{(u)} b_{jkr}^{(u)} (t_{k+iu} - t_k)^{r_0} (t_{k+ju} - t_k)^{r_0} \\ \times \iint_{[0,1]^2} \frac{((1-v)(1-w))^{r_0-1}}{((r_0-1)!)^2} |\dot{v}_{ik} - \dot{w}_{jk}|^{2\beta_0} d_0(t_k) dv dw. \end{aligned} \quad (6.13)$$

From the equivalents (6.3) and (6.7), we can write the leading term of (6.13) as a Riemann sum on t_k to obtain

$$\begin{aligned} \delta_n^{2p-2\beta_0} \mathbb{E}(\overline{(D_r^{(u)} X)^2}) \xrightarrow{n \rightarrow \infty} -\frac{1}{2} \left(\frac{u}{T}\right)^{-2p+2\beta_0} \int_0^T d_0(t) \psi^{2p+1-2\beta_0}(t) dt \\ \times \sum_{i,j=0}^r \frac{(ij)^{r_0}}{\prod_{\substack{m=0 \\ m \neq i}}^r (i-m) \prod_{\substack{q=0 \\ q \neq j}}^r (j-q)} \iint_{[0,1]^2} \frac{((1-v)(1-w))^{r_0-1}}{((r_0-1)!)^2} |iv - jw|^{2\beta_0} dv dw. \end{aligned}$$

Next by performing elementary but tedious multiple integrations by parts, we arrive at the following simpler form of $\ell(r, r_0, \beta_0)$ given in (6.8), for $n\delta_n \rightarrow T$.

Case $r_0 = 0$, $r = r_0 + 1$, $r_0 + 2$. The proof is the same but starting from (6.10) and $\ell = k$.

Case $r_0 \geq 1$ and $r = 1, \dots, r_0$. In this case, $r^* = r$ and from the relation (6.11), one gets

$$\begin{aligned} \mathbb{E}(D_{r,k}^{(u)} X)^2 &= \sum_{i,j=0}^r b_{ikr}^{(u)} b_{jkr}^{(u)} (t_{k+iu} - t_k)^r (t_{k+ju} - t_k)^r \\ &\quad \times \iint_{[0,1]^2} \frac{((1-v)(1-w))^{r-1}}{((r-1)!)^2} \mathbb{L}^{(r,r)}(\dot{v}_{ik}, \dot{w}_{jk}) dv dw. \end{aligned}$$

The result follows after Riemann summation with the help of uniform continuity of $\mathbb{L}^{(r,r)}(\cdot, \cdot)$, $r = 1, \dots, r_0$ and properties (6.3), (6.5). \square

6.2. Auxiliary results. The following lemma gives some useful results on the asymptotic behavior of $\mathbb{C}_r(k, \ell)$ and $\mathbb{C}_r^2(k, \ell)$ with $\mathbb{C}_r(k, \ell) = \text{Cov}(D_{r,k}^{(u)} X, D_{r,\ell}^{(u)} X)$ with $n_r = n - ur$ and u a positive integer.

Lemma 6.4. *Suppose that Assumption A2.1 and A2.2 are fulfilled.*

(i) *Under the condition A2.1-(iii-1) and for $r = r_0 + p$, $p = 1$ or $p = 2$, one obtains*

$$\max_{k=0, \dots, n_r} \sum_{\ell=0}^{n_r} |\mathbb{C}_r(k, \ell)| = \begin{cases} \mathcal{O}(n^{2p-2\beta_0}) & \text{if } 0 < \beta_0 < \frac{1}{2}, \\ \mathcal{O}(n^{2p-1} \ln n) & \text{if } \beta_0 = \frac{1}{2}, \\ \mathcal{O}(n^{2p-1}) & \text{if } \frac{1}{2} < \beta_0 < 1; \end{cases}$$

and

$$\sum_{k=0}^{n_r} \sum_{\ell=0}^{n_r} \mathbb{C}_r^2(k, \ell) = \begin{cases} \mathcal{O}(n^{4p-4\beta_0+1}) & \text{if } 0 < \beta_0 < \frac{3}{4}, \\ \mathcal{O}(n^{4p-2} \ln n) & \text{if } \beta_0 = \frac{3}{4}, \\ \mathcal{O}(n^{4p-2}) & \text{if } \frac{3}{4} < \beta_0 < 1. \end{cases}$$

(ii) Under the condition A2.1-(iii-2) and for $r = r_0 + 2$, one obtains

$$\max_{k=0, \dots, n_r} \sum_{\ell=0}^{n_r} |\mathbb{C}_r(k, \ell)| = \mathcal{O}(n^{4-2\beta_0}) \text{ and } \sum_{k=0}^{n_r} \sum_{\ell=0}^{n_r} \mathbb{C}_r^2(k, \ell) = \mathcal{O}(n^{9-4\beta_0}).$$

(iii) If $r = 1, \dots, r_0$ (with $r_0 \geq 1$), then $\max_{k=0, \dots, n_r} \sum_{\ell=0}^{n_r} |\mathbb{C}_r(k, \ell)| = \mathcal{O}(n)$ and

$$\sum_{k=0}^{n_r} \sum_{\ell=0}^{n_r} \mathbb{C}_r^2(k, \ell) = \mathcal{O}(n^2).$$

Proof. (i) Setting $\mu(t) = \mathbb{E}(X(t))$, μ is r_0 -times differentiable and similarly to (6.10)-(6.11), we get the expansion

$$\begin{aligned} \mathbb{C}_r(k, \ell) &= \sum_{i,j=0}^r b_{ikr}^{(u)} b_{j\ell r}^{(u)} (t_{k+iu} - t_k)^{r_0} (t_{\ell+ju} - t_\ell)^{r_0} \\ &\quad \times \iint_{[0,1]^2} \frac{((1-v)(1-w))^{r_0-1}}{((r_0-1)!)^2} \mathbb{K}^{(r_0, r_0)}(\dot{v}_{ik}, \dot{w}_{j\ell}) dv dw \end{aligned}$$

for $r_0 \geq 1$ while if $r_0 = 0$, $\mathbb{C}_r(k, \ell) = \sum_{i=0}^r \sum_{j=0}^r b_{ikr}^{(u)} b_{j\ell r}^{(u)} \mathbb{K}(t_{k+iu}, t_{\ell+ju})$.

Case $r = r_0 + 1$ or $r_0 + 2$. For $r_0 \geq 1$, we have the bound:

$$\max_{k=0, \dots, n_r} \sum_{\ell=0}^{n_r} |\mathbb{C}_r(k, \ell)| \leq U_{1n} + U_{2n} + U_{3n}$$

with $U_{1n} = \max_{k=ur+1, \dots, n_r} \sum_{\ell=0}^{k-ur-1} |\mathbb{C}_r(k, \ell)|$, $U_{2n} = \max_{k=0, \dots, n-2ur-1} \sum_{\ell=k+ur+1}^{n_r} |\mathbb{C}_r(k, \ell)|$

and $U_{3n} = \max_{k=0, \dots, n_r} \sum_{\ell=\max(0, k-ur)}^{\min(n_r, k+ur)} |\mathbb{C}_r(k, \ell)|$. First, consider the sum $U_{1n} + U_{2n}$ where

$|k - \ell| \geq ur + 1$. Since $\sum_{i=0}^r b_{ikr}^{(u)} (t_{k+iu} - t_k)^{r_0} = 0$ for $r = r_0 + 1$ or $r = r_0 + 2$, and $[t_k, \dot{v}_{ik}]$ is distinct from $[t_\ell, \dot{w}_{j\ell}]$, we get

$$\begin{aligned} \mathbb{C}_r(k, \ell) &= \sum_{i,j=0}^r b_{ikr}^{(u)} b_{j\ell r}^{(u)} (t_{k+iu} - t_k)^{r_0} (t_{\ell+ju} - t_\ell)^{r_0} \\ &\quad \times \iint_{[0,1]^2} \frac{((1-v)(1-w))^{r_0-1}}{((r_0-1)!)^2} \int_{t_k}^{\dot{v}_{ik}} \int_{t_\ell}^{\dot{w}_{j\ell}} \mathbb{K}^{(r_0+1, r_0+1)}(s, t) ds dt dv dw. \end{aligned} \quad (6.14)$$

Condition A2.1-(iii-1), together with the bounds (6.2) and (6.6), gives a bound of $\mathcal{O}(n^{2p-2\beta_0} \sum_{i=1}^n i^{-2(1-\beta_0)})$ for $|U_{1n} + U_{2n}|$, which is of order $n^{2(p-\beta_0)}$ if $0 < \beta_0 < \frac{1}{2}$, $n^{2(p-\beta_0)} \ln n$ if $\beta_0 = \frac{1}{2}$ and n^{2p-1} if $\beta_0 > \frac{1}{2}$. Next, for U_{3n} where $|k - \ell| \leq ur$, we obtain that $U_{3n} = \mathcal{O}(n^{2(p-\beta_0)})$ in a similar way as in the proof of Proposition 6.1, and with the help of Cauchy-Schwarz inequality to control the terms depending on $\mu^{(r_0)}(t)$.

We proceed similarly for the case $r_0 = 0$, starting from the definition of $\mathbb{C}_r(k, \ell)$ as well as for the study of $\sum_{k=0}^{n_r} \sum_{\ell=0}^{n_r} \mathbb{C}_r^2(k, \ell)$ for which dominant terms are of order $\mathcal{O}(n^{1+4p-4\beta_0} \sum_{i=1}^n i^{-4(1-\beta_0)})$.

(ii) The condition A2.1-(iii-2) and $r = r_0 + 2$ allows to transform (6.14) into

$$\begin{aligned} \mathbb{C}_r(k, \ell) &= \sum_{i,j=0}^r b_{ikr}^{(u)} b_{jkr}^{(u)} (t_{k+iu} - t_k)^{r_0} (t_{\ell+ju} - t_\ell)^{r_0} \iint_{[0,1]^2} \frac{((1-v)(1-w))^{r_0-1}}{((r_0-1)!)^2} \\ &\quad \times \int_{t_k}^{v_{ik}} \int_{t_\ell}^{w_{j\ell}} \int_{t_k}^t \int_{t_\ell}^s K^{(r_0+2, r_0+2)}(y, z) dy dz ds dt dv dw \end{aligned} \quad (6.15)$$

which gives that

$$\max_{k=0, \dots, n_r} \sum_{\ell=0}^{n_r} |\mathbb{C}_r(k, \ell)| = \mathcal{O}(n^{2(2-\beta_0)} \sum_{i=1}^n i^{-4+2\beta_0}) = \mathcal{O}(n^{2(2-\beta_0)})$$

for all $\beta_0 \in]0, 1[$ and $r_0 \geq 1$. From (6.15), we also get that $\sum_{k=0}^{n_r} \sum_{\ell=0}^{n_r} \mathbb{C}_r^2(k, \ell) = \mathcal{O}(n^{9-4\beta_0} \sum_{i=1}^n i^{-8+4\beta_0}) = \mathcal{O}(n^{9-4\beta_0})$ for all $\beta_0 \in]0, 1[$.

(iii) Results of these part, where $r_0 \geq 1$, are consequences of

$$\begin{aligned} \mathbb{C}_r(k, \ell) &= \sum_{i,j=0}^r b_{ikr}^{(u)} b_{j\ell r}^{(u)} (t_{k+iu} - t_k)^r (t_{\ell+ju} - t_\ell)^r \\ &\quad \times \iint_{[0,1]^2} \frac{((1-v)(1-w))^{r-1}}{(r-1)!^2} K^{(r,r)}(v_{ik}, w_{j\ell}) dv dw = \mathcal{O}(1) \end{aligned}$$

with uniform continuity of $K^{(r,r)}(\cdot, \cdot)$ for $r = 1, \dots, r_0$ together with bounds (6.2) and (6.6). \square

Next proposition gives a general exponential bound, involved in all our results.

Proposition 6.2. *Suppose that Assumption A2.1 and A2.2 are fulfilled. Let $\eta_n(r)$ be some given positive sequence and $u \in \mathbb{N}^*$, then*

$$\mathbb{P}\left(\left|\overline{(D_r^{(u)} X)^2} - \mathbb{E}\left(\overline{(D_r^{(u)} X)^2}\right)\right| \geq \eta_n(r)\right)$$

is of order:

$$\begin{aligned} \mathcal{O}\left(\exp\left(-C(r)n\eta_n(r) \times \min\left(\left(\max_{0 \leq k \leq n_r} \sum_{\ell=0}^{n_r} |\mathbb{C}_r(k, \ell)|\right)^{-1}, \frac{n\eta_n(r)}{\sum_{k,\ell=0}^{n_r} \mathbb{C}_r^2(k, \ell)}\right)\right)\right) \\ + \mathcal{O}\left(\frac{v_n^{1/2}(r)}{n\eta_n(r)} \exp\left(-C(r)\frac{n^2\eta_n^2(r)}{v_n(r)}\right)\right) \end{aligned}$$

for some positive constant $C(r)$, not depending on $\eta_n(r)$ and

$$v_n(r) := n \max_{k=0, \dots, n_r} (\mathbb{E}(D_{r,k}^{(u)} X))^2 \max_{k=0, \dots, n_r} \sum_{\ell=0}^{n_r} |\mathbb{C}_r(k, \ell)|. \quad (6.16)$$

Proof. For all $r \geq 1$, we may bound $\mathbb{P}\left(\left|\overline{(D_r^{(u)} X)^2} - \mathbb{E}\left(\overline{(D_r^{(u)} X)^2}\right)\right| \geq \eta_n(r)\right)$ by $S_1 + S_2$ with

$$S_1 = \mathbb{P}\left(\left|\sum_{k=0}^{n_r} (D_{r,k}^{(u)} X - \mathbb{E}(D_{r,k}^{(u)} X))^2 - \text{Var}(D_{r,k}^{(u)} X)\right| > \frac{(n_r+1)\eta_n(r)}{2}\right)$$

and $S_2 = \mathbb{P}\left(\left|\sum_{k=0}^{n_r} (\mathbb{E}(D_{r,k}^{(u)} X))(D_{r,k}^{(u)} X - \mathbb{E}(D_{r,k}^{(u)} X))\right| > \frac{(n_r+1)\eta_n(r)}{4}\right)$. First, let $\{Y_i\}_{i=1, \dots, d_n}$ be an orthonormal basis for the linear span of $\{D_{r,k}^{(u)} X\}_{k=0, \dots, n_r}$ (so

that Y_i are i.i.d. with density $\mathcal{N}(0, 1)$. We can write $D_{r,k}^{(u)} X - \mathbb{E}(D_{r,k}^{(u)} X) = \sum_{i=1}^{d_n} d_{k,i} Y_i$ with $d_{k,i} = \text{Cov}(D_{r,k}^{(u)} X, Y_i)$. Next, if $Y = (Y_1, \dots, Y_{d_n})^\top$, we obtain

$$\sum_{k=0}^{n_r} (D_{r,k}^{(u)} X - \mathbb{E} D_{r,k}^{(u)} X)^2 = \sum_{i,j=1}^{d_n} c_{i,j} Y_i Y_j = Y^\top C Y \text{ and } \sum_{k=0}^{n_r} \text{Var}(D_{r,k}^{(u)} X) = \sum_{i=1}^{d_n} c_{i,i}$$

with $c_{i,j} = \sum_{k=0}^{n_r} d_{k,i} d_{k,j}$. Next, for $C = (c_{i,j})_{i,j=1,\dots,d_n}$ and $D = (d_{k,j})_{k=0,\dots,n_r, j=1,\dots,d_n}$, one gets $C = D^\top D$ where C is a real, symmetric and positive semidefinite matrix. There exists an orthogonal matrix P such that $\text{diag}(\lambda_1, \dots, \lambda_{d_n}) = P^\top C P$, for λ_i eigenvalues of C . Then we can transform the quadratic form as:

$$\sum_{k=0}^{n_r} (D_{r,k}^{(u)} X - \mathbb{E}(D_{r,k}^{(u)} X))^2 = \sum_{i=1}^{d_n} \lambda_i (P^\top Y)_i^2$$

where $(P^\top Y)_i$ denotes the i -th component of the $(d_n \times 1)$ vector $P^\top Y$. Since $\sum_{i=1}^{d_n} c_{i,i} = \sum_{i=1}^{d_n} \lambda_i$, we arrive at

$$S_1 = \mathbb{P}\left(\left|\sum_{i=1}^{d_n} \lambda_i ((P^\top Y)_i^2 - 1)\right| \geq \frac{(n_r + 1)\eta_n(r)}{2}\right).$$

Now, with the exponential bound of Hanson and Wright (1971), we obtain for some generic constant c :

$$S_1 \leq 2 \exp\left(-c(n_r + 1)\eta_n(r) \times \min\left(\frac{1}{\max(\lambda_i)}, \frac{(n_r + 1)\eta_n(r)}{\sum \lambda_i^2}\right)\right).$$

Next, since $D^\top D$ and DD^\top have the same non zero eigenvalues,

$$\max_{i=1,\dots,d_n} \lambda_i \leq \max_{0 \leq k \leq n_r} \sum_{\ell=0}^{n_r} |\mathbb{C}_r(k, \ell)|$$

and $\sum_{i=1}^{d_n} \lambda_i^2 = \sum_{i=1}^{d_n} \sum_{j=1}^{d_n} c_{i,j} c_{j,i} = \sum_{k=0}^{n_r} \sum_{\ell=0}^{n_r} (\sum_{i=1}^{d_n} d_{k,i} d_{\ell,i})^2 = \sum_{k=0}^{n_r} \sum_{\ell=0}^{n_r} \mathbb{C}_r^2(k, \ell)$. Finally S_1 is bounded by

$$2 \exp\left(-c(n_r + 1)\eta_n(r) \times \min\left(\left(\max_{0 \leq k \leq n_r} \sum_{\ell=0}^{n_r} |\mathbb{C}_r(k, \ell)|\right)^{-1}, \frac{(n_r + 1)\eta_n(r)}{\sum_{k=0}^{n_r} \sum_{\ell=0}^{n_r} \mathbb{C}_r^2(k, \ell)}\right)\right).$$

For S_2 , we use the classical exponential bound on a Gaussian variable: $Y \sim \mathcal{N}(0, \sigma^2)$ implies that $\mathbb{P}(|Y| \geq \varepsilon) \leq \min(1, \sqrt{\frac{2\sigma^2}{\pi\varepsilon^2}} \exp(-\frac{\varepsilon^2}{2\sigma^2}))$, $\varepsilon > 0$. Here $Y = \sum_{k=0}^{n_r} (\mathbb{E} D_{r,k}^{(u)} X)(D_{r,k}^{(u)} X - \mathbb{E} D_{r,k}^{(u)} X)$ and we get easily that $\text{Var}(Y) \leq v_n(r)$. \square

6.3. Proofs of section 3.

Proof. Proof of Theorem 3.1

Recall that \hat{r}_0 is given by: $\hat{r}_0 = \min\left\{r \in \{2, \dots, m_n\} : B_n(r) \text{ holds}\right\} - 2$

where the event $B_n(r)$ is defined by $B_n(r) = \{(D_r^{(1)} X)^2 \geq n^2 b_n\}$, and $\hat{r}_0 = \ell_0$ if $\cap_{r=2}^{m_n} B_n(r)$. The condition $m_n \rightarrow \infty$ guarantees that for n large enough, $r_0 + 2 \in \{2, \dots, m_n\}$. From this definition, we write

$$\mathbb{E}(\hat{r}_0 - r_0)^2 = \sum_{r=0}^{m_n-2} (r - r_0)^2 \mathbb{P}(\hat{r}_0 = r) + (\ell_0 - r_0)^2 \mathbb{P}(\hat{r}_0 = \ell_0)$$

where $\mathbb{P}(\widehat{r}_0 = 0) = \mathbb{P}(B_n(2))$, $\mathbb{P}(\widehat{r}_0 = r) = \mathbb{P}(B_n^c(2) \cap \dots \cap B_n^c(r+1) \cap B_n(r+2))$ if $r = 1, \dots, m_n - 2$, and $\mathbb{P}(\widehat{r}_0 = l_0) \leq \mathbb{P}(B_n^c(r_0 + 2))$. Then, for all $r_0 \in \mathbb{N}_0$: $\mathbb{E}(\widehat{r}_0 - r_0)^2 = \mathcal{O}(T_{1n}(r_0)) + \mathcal{O}(m_n^3 T_{2n}(r_0))$ where we have set $T_{1n}(0) = 0$, $T_{1n}(r_0) = \sum_{r=2}^{r_0+1} \mathbb{P}(B_n(r))$ (for $r_0 \geq 1$) and $T_{2n}(r_0) = \mathbb{P}(B_n^c(r_0 + 2))$. Now, the study of T_{1n} and T_{2n} is derived from results of Lemma 6.2, Lemma 6.3, Proposition 6.1 and Lemma 6.4. In particular, since $\mu \in C^{r_0+1}([0, T])$ we get:

$$\mathbb{E}(D_{r,k}^{(u)} X) = \sum_{i=0}^r b_{ikr}(u)(t_{k+iu} - t_k)^{r^*} \int_0^1 \frac{(1-v)^{r^*-1}}{(r^*-1)!} \mu^{(r^*)}(t_k + (t_{k+iu} - t_k)v) dv$$

which is $\mathcal{O}(n^{r-r^*})$ for $r^* = \min(r, r_0 + 1)$ implying that $\mathbb{E}(D_{r,k}^{(u)} X) = \mathcal{O}(1)$ for $r = 1, \dots, r_0 + 1$, $\mathbb{E}(D_{r,k}^{(u)} X) = \mathcal{O}(n)$ for $r = r_0 + 2$. Then one may bound $v_n(r)$ given in equation (6.16) by $\mathcal{O}(n^2)$ if $r = 1, \dots, r_0$, $\mathcal{O}(n^{3-2\beta_0} \mathbb{1}_{]0, \frac{1}{2}[}(\beta_0) + n^2 \ln n \mathbb{1}_{\{\frac{1}{2}\}}(\beta_0) + n^2 \mathbb{1}_{] \frac{1}{2}, 1[}(\beta_0))$ if $r = r_0 + 1$ with A2.1-(iii-1), $\mathcal{O}(n^{7-2\beta_0} \mathbb{1}_{]0, \frac{1}{2}[}(\beta_0) + n^6 \ln n \mathbb{1}_{\{\frac{1}{2}\}}(\beta_0) + n^6 \mathbb{1}_{] \frac{1}{2}, 1[}(\beta_0))$ if $r = r_0 + 2$ with A2.1-(iii-1), and $\mathcal{O}(n^{7-2\beta_0})$ if $r = r_0 + 2$ and A2.1-(iii-2) holds. Next after some calculations based on properties $n^{2\beta_0} b_n \rightarrow \infty$ and $n^{-2(1-\beta_0)} b_n \rightarrow 0$, one may derive from Proposition 6.2 that:

$$T_{1n}(r_0) = \mathcal{O}\left(\exp\left(-D(r_0)b_n(n^{2\beta_0+1} \mathbb{1}_{]0, \frac{1}{2}[}(\beta_0) + \left(\frac{n^2}{\ln n}\right) \mathbb{1}_{\{\frac{1}{2}\}}(\beta_0) + n^2 \mathbb{1}_{] \frac{1}{2}, 1[}(\beta_0))\right)\right).$$

Next, if A2.1-(iii-1) holds

$$T_{2n}(r_0) = \mathcal{O}\left(\exp\left(-D(r_0)(n \mathbb{1}_{]0, \frac{1}{2}[}(\beta_0) + \left(\frac{n}{\ln n}\right) \mathbb{1}_{\{\frac{1}{2}\}}(\beta_0) + n^{2(1-\beta_0)} \mathbb{1}_{] \frac{1}{2}, 1[}(\beta_0))\right)\right)$$

while, under A2.1-(iii-2) and for all $\beta_0 \in]0, 1[$, $T_{2n}(r_0) = \mathcal{O}(\exp(-D(r_0)n))$. For $p = 1, 2$, we get that $T_{1n}(r_0) = o(T_{2n}(r_0))$ and the mean square error follows. Finally, to obtain a bound for $\mathbb{P}(\widehat{r}_0 \neq r_0)$, it suffices to notice that $\{\widehat{r}_0 = 0\} = B_n(2)$ for $r_0 = 0$ and $\{\widehat{r}_0 = r_0\} = B_n^c(2) \cap \dots \cap B_n^c(r_0 + 1) \cap B_n(r_0 + 2)$ for $r_0 \geq 1$, by this way $\mathbb{P}(\widehat{r}_0 \neq r_0) = T_{1n}(r_0) + T_{2n}(r_0) = T_{2n}(r_0)(1 + o(1))$. \square

Proof. Proof of Theorem 3.2

We start the proof, with either $p = 1$ or $p = 2$, and thus denote by \widehat{r}_p (resp. r_p) the quantity $\widehat{r}_0 + p$ (resp. $r_0 + p$). We set

$$l_n(p, r_0, \beta_0) = -\frac{1}{2n} \sum_{k=0}^n d_0(t_k) \psi^{2(p-\beta_0)}(t_k) \sum_{i,j=0}^{r_p} \frac{(ij)^{r_0}}{\prod_{\substack{m=0 \\ m \neq i}}^{r_p} (i-m) \prod_{\substack{q=0 \\ q \neq j}}^{r_p} (j-q)} \\ \times \iint_{[0,1]^2} \frac{((1-v)(1-w))^{r_0-1}}{(r_0-1)!^2} |iv - jw|^{2\beta_0} dv dw, \quad (6.17)$$

for all $r_0 \geq 1$ while if $r_0 = 0$,

$$l_n(p, 0, \beta_0) = -\frac{1}{2n} \sum_{k=0}^n d_0(t_k) \psi^{2(p-\beta_0)}(t_k) \sum_{i,j=0}^{r_p} \frac{|i-j|^{2\beta_0}}{\prod_{\substack{m=0 \\ m \neq i}}^p (i-m) \prod_{\substack{q=0 \\ q \neq j}}^p (j-q)}. \quad (6.18)$$

We study the convergence of $\widehat{\alpha}_p = 2(\widehat{\beta}_n^{(p)} - p)$ toward $\alpha_p = 2(\beta_0 - p)$, so that

$$\widehat{\alpha}_p = \frac{\ln((D_{\widehat{r}_p}^{(u)} X)^2) - \ln((D_{\widehat{r}_p}^{(v)} X)^2)}{\ln(u/v)}.$$

We consider the following decomposition of $\ln(u/v)\hat{\alpha}_p$:

$$\begin{aligned} & \ln\left(\frac{n^{\alpha_p}}{n - u\hat{r}_p + 1} \sum_{k=0}^{n-u\hat{r}_p} (D_{\hat{r}_p, k}^{(u)} X)^2 - u^{\alpha_p} l_n(p, r_0, \beta_0) + u^{\alpha_p} l_n(p, r_0, \beta_0)\right) \\ & - \ln\left(\frac{n^{\alpha_p}}{n - v\hat{r}_p + 1} \sum_{k=0}^{n-v\hat{r}_p} (D_{\hat{r}_p, k}^{(v)} X)^2 - v^{\alpha_p} l_n(p, r_0, \beta_0) + v^{\alpha_p} l_n(p, r_0, \beta_0)\right) \end{aligned}$$

Hence $\ln(u/v)(\hat{\alpha}_p - \alpha_p) = F_n(u) - F_n(v) + o(F_n(u) + F_n(v))$ where $o(\cdot) \xrightarrow[n \rightarrow \infty]{a.s.} 0$ as soon as $F_n(\cdot) \xrightarrow[n \rightarrow \infty]{a.s.} 0$ with

$$F_n(u) = \frac{n^{\alpha_p} \overline{(D_{\hat{r}_p}^{(u)} X)^2} - u^{\alpha_p} l_n(p, r_0, \beta_0)}{u^{\alpha_p} l_n(p, r_0, \beta_0)} = \frac{F_{1,n,p}(u) + F_{2,n,p}(u) + F_{3,n,p}(u)}{u^{\alpha_p} l_n(p, r_0, \beta_0)}$$

for $F_{1,n,p}(u) = n^{\alpha_p} \left(\overline{(D_{\hat{r}_p}^{(u)} X)^2} - \overline{(D_{r_p}^{(u)} X)^2} \right)$, $F_{2,n,p}(u) = n^{\alpha_p} \left(\overline{(D_{\hat{r}_p}^{(u)} X)^2} - \mathbb{E} \overline{(D_{r_p}^{(u)} X)^2} \right)$

and $F_{3,n,p}(u) = n^{\alpha_p} \mathbb{E} \left(\overline{(D_{\hat{r}_p}^{(u)} X)^2} \right) - u^{\alpha_p} l_n(p, r_0, \beta_0)$.

(i) **Study of $F_{1,n,p}(u)$.** From Theorem 3.1, we get that $\sum_n \mathbb{P}(\hat{r}_0 \neq r_0) < \infty$, so, a.s.

for n large enough, $\hat{r}_0 = r_0$ and $F_{1,n,p}(u) \equiv 0$, $p = 1$ or $p = 2$.

(ii) **Study of $F_{2,n,p}(u)$.** We study

$$\mathbb{P} \left(\left| \overline{(D_{\hat{r}_p}^{(u)} X)^2} - \mathbb{E} \overline{(D_{r_p}^{(u)} X)^2} \right| > c_p n^{2(p-\beta_0)} \psi_{np}^{-1}(\beta_0) \right)$$

for c_p a positive constant, $\psi_{n2}(\beta_0) \equiv \left(\frac{n}{\ln n} \right)^{\frac{1}{2}}$ and

$$\psi_{n1}(\beta_0) = \left(\frac{n}{\ln n} \right)^{\frac{1}{2}} \mathbb{1}_{[0, \frac{3}{4}]}(\beta_0) + \left(\frac{n^{1/2}}{\ln n} \right) \mathbb{1}_{\{\frac{3}{4}\}}(\beta_0) + \left(\frac{n^{2(1-\beta_0)}}{\ln n} \right) \mathbb{1}_{] \frac{3}{4}, 1]}(\beta_0).$$

We apply Lemma 6.4 and Proposition 6.2 with $p = 1$ or $p = 2$. After some calculations and the application of Borel Cantelli's lemma with c_p chosen large enough, we obtain that for $p = 1$, almost surely, $\lim_{n \rightarrow \infty} \psi_{np}(\beta_0) |F_{2,n,p}(u)| < +\infty$ under the condition A2.1-(iii-p), where $p = 1$ or 2 .

(iii) **Study of $F_{3,n,p}(u)$.** From (6.11) and proceeding similarly as in (6.12), we get for $r_0 \geq 1$, that $n^{\beta_1} (n^{\alpha_p} \mathbb{E} \overline{(D_{\hat{r}_p}^{(u)} X)^2} - u^{\alpha_p} l_n(p, r_0, \beta_0))$ could be decomposed into $B_{n1} + B_{n2} + B_{n3}$ with

$$\begin{aligned} B_{n1} = & -\frac{n^{\alpha_p+\beta_1}}{2(n - ur_p + 1)} \sum_{k=0}^{n-ur_p} \sum_{i,j=0}^{r_p} b_{ikr}^{(u)} b_{jkr}^{(u)} (t_{k+iu} - t_k)^{r_0} (t_{k+ju} - t_k)^{r_0} \\ & \times \iint_{[0,1]^2} \frac{((1-v)(1-w))^{r_0-1}}{(r_0-1)!^2} |\dot{v}_{ik} - \dot{w}_{jk}|^{2\beta_0+\beta_1} \\ & \times \left\{ \frac{\mathbb{L}^{(r_0, r_0)}(\dot{v}_{ik}, \dot{w}_{jk}) - \frac{1}{2} \mathbb{L}^{(r_0, r_0)}(\dot{v}_{ik}, \dot{v}_{ik}) - \frac{1}{2} \mathbb{L}^{(r_0, r_0)}(\dot{w}_{jk}, \dot{w}_{jk})}{|\dot{v}_{ik} - \dot{w}_{jk}|^{2\beta_0}} - d_0(\dot{w}_{jk}) \right. \\ & \left. - \frac{d_0(\dot{w}_{jk})}{|\dot{v}_{ik} - \dot{w}_{jk}|^{\beta_1}} - d_1(\dot{w}_{jk}) \right\} dv dw \end{aligned}$$

$$\begin{aligned} B_{n2} = & -\frac{n^{\alpha_p+\beta_1}}{2(n - ur_p + 1)} \sum_{k=0}^{n-ur_p} \sum_{i,j=0}^{r_p} b_{ikr}^{(u)} b_{jkr}^{(u)} (t_{k+iu} - t_k)^{r_0} (t_{k+ju} - t_k)^{r_0} \\ & \times \iint_{[0,1]^2} \frac{((1-v)(1-w))^{r_0-1}}{(r_0-1)!^2} |\dot{v}_{ik} - \dot{w}_{jk}|^{2\beta_0+\beta_1} d_1(\dot{w}_{jk}), \end{aligned}$$

$$B_{n3} = n^{\beta_1} \left(\frac{-n^{\alpha_p}}{2(n - ur_p + 1)} \sum_{k=0}^{n-ur_p} \sum_{i,j=0}^{r_p} b_{ikr}^{(u)} b_{jkr}^{(u)} (t_{k+iu} - t_k)^{r_0} (t_{k+ju} - t_k)^{r_0} \right. \\ \left. \times \iint_{[0,1]^2} \frac{((1-v)(1-w))^{r_0-1}}{(r_0-1)!^2} |\dot{v}_{ik} - \dot{w}_{jk}|^{2\beta_0} d_0(\dot{w}_{jk}) dv dw - u^{\alpha_p} l_n(p, r_0, \beta_0) \right)$$

with $l_n(p, r_0, \beta_0)$ given by (6.17). Next, using Lemma 6.2 and 6.3 and the condition (3.1) with uniform continuity of $d_1(\cdot)$, we get that $B_{n1} = o(1)$ and B_{n2} has the limit:

$$- \frac{u^{\alpha_p + \beta_1}}{2} \sum_{i,j=0}^{r_p} \frac{(ij)^{r_0} \int_0^T d_1(t) \psi^{1-\alpha_p-\beta_1}(t) dt}{\prod_{\substack{m=0 \\ m \neq i}}^{r_p} (i-m) \prod_{\substack{q=0 \\ q \neq j}}^{r_p} (j-q)} \\ \times \iint_{[0,1]^2} \frac{((1-v)(1-w))^{r_0-1}}{((r_0-1)!)^2} |iv - jw|^{2\beta_0 + \beta_1} dv dw.$$

For the last term B_{n3} , one may show that it is of order $\mathcal{O}(n^{\beta_1-1})$. Finally, the case $r_0 = 0$ is treated similarly from (6.10).

Conclusion. One may note that the determinist term, $l_n(p, r_0, \beta_0)$, defined in (6.17)-(6.18), converges to the nonzero term:

$$- \frac{1}{2} \sum_{i,j=0}^{r_p} \frac{(ij)^{r_0} \int_0^T d_0(t) \psi^{-\alpha_p+1}(t) dt}{\prod_{\substack{m=0 \\ m \neq i}}^{r_p} (i-m) \prod_{\substack{q=0 \\ q \neq j}}^{r_p} (j-q)} \iint_{[0,1]^2} \frac{((1-v)(1-w))^{r_0-1}}{(r_0-1)!^2} |iv - jw|^{2\beta_0} dv dw$$

for $r_0 \geq 1$ while if $r_0 = 0$, the limit is $-\frac{1}{2} \sum_{i,j=0}^p \frac{|i-j|^{2\beta_0} \int_0^T d_0(t) \psi^{-\alpha_p+1}(t) dt}{\prod_{\substack{m=0 \\ m \neq i}}^p (i-m) \prod_{\substack{q=0 \\ q \neq j}}^p (j-q)}$. \square

6.4. Proofs of section 4.

Proof. Proof of Theorem 4.1

We set $\tilde{r}_0 = \max(\hat{r}_0, 1)$ and, for \hat{r}_0 and $\tilde{X}_r(\cdot)$ respectively defined in (2.5) and (4.1), we use the convention: $\tilde{X}_{\tilde{r}_0}(\cdot) = \tilde{X}_{m_n-1}(\cdot)$ and $\tilde{X}_{\tilde{r}_0+1}(\cdot) = \tilde{X}_{m_n}(\cdot)$ when $\hat{r}_0 = l_0$.

(a) If $\bar{r} = \max(r, 1)$ and $\bar{r}_0 = \max(r_0, 1)$, we get, for n large enough such that $r_0 \leq m_n - 2$,

$$(X(t) - \tilde{X}_{\tilde{r}_0}(t))^2 = \sum_{r=0}^{m_n-2} (X(t) - \tilde{X}_{\bar{r}}(t))^2 \mathbb{1}_{\{\hat{r}_0=r\}} + (X(t) - \tilde{X}_{m_n-1}(t))^2 \mathbb{1}_{\{\hat{r}_0=l_0\}} \\ \leq (X(t) - \tilde{X}_{\bar{r}_0}(t))^2 + \mathbb{1}_{\{\hat{r}_0 \neq r_0\}} \sum_{r=0, r \neq r_0}^{m_n-1} (X(t) - \tilde{X}_{\bar{r}}(t))^2.$$

By this way, $e_p^2(\text{app}(\hat{r}_0))$ should be bounded by

$$\int_0^T \mathbb{E} (X(t) - \tilde{X}_{\bar{r}_0}(t))^2 \rho(t) dt + (\mathbb{P}(\hat{r}_0 \neq r_0))^{\frac{1}{2}} \sum_{r=0, r \neq r_0}^{m_n-1} \int_0^T \left(\mathbb{E} (X(t) - \tilde{X}_{\bar{r}}(t))^4 \right)^{\frac{1}{2}} \rho(t) dt$$

We make use of the exponential bound established for $\mathbb{P}(\hat{r}_0 \neq r_0)$ in Theorem 3.1 as well as the property $\mathbb{E}(Y^4) \leq 3(\mathbb{E}(Y^2))^2$ for a Gaussian r.v. Y . Moreover,

$\sup_{t \in [0, T]} \left(\mathbb{E} (X(t) - \tilde{X}_r(t))^2 \right) = \max_{k=0, \dots, \lfloor \frac{T}{\tau} \rfloor - 1} \sup_{t \in \mathcal{I}_k} \left(\mathbb{E} (X(t) - \tilde{X}_r(t))^2 \right)$. If $r_0 \geq 1$, we use the decomposition established in Blanke and Vial (2008, lemma 4.1) to obtain,

for $t \in \mathcal{I}_k$ and $r^* = \min(r, r_0)$:

$$\begin{aligned} \mathbb{E}(X(t) - \tilde{X}_r(t))^2 &= \sum_{i,j=0}^r L_{i,k,r}(t) L_{j,k,r}(t) \frac{(t_{kr+i} - t_{kr})^{r^*} (t_{kr+j} - t_{kr})^{r^*}}{((r^* - 1)!)^2} \\ &\quad \times \iint_{[0,1]^2} ((1-v)(1-w))^{r^*-1} \left\{ \mathbb{L}^{(r^*, r^*)}(t_{kr} + (t - t_{kr})v, t_{kr} + (t - t_{kr})w) \right. \\ &\quad - \mathbb{L}^{(r^*, r^*)}(t_{kr} + (t - t_{kr})v, t_{kr} + (t_{kr+j} - t_{kr})w) \\ &\quad - \mathbb{L}^{(r^*, r^*)}(t_{kr} + (t_{kr+i} - t_{kr})v, t_{kr} + (t - t_{kr})w) \\ &\quad \left. + \mathbb{L}^{(r^*, r^*)}(t_{kr} + (t_{kr+i} - t_{kr})v, t_{kr} + (t_{kr+j} - t_{kr})w) \right\} dv dw. \end{aligned}$$

If $r = 1, \dots, r_0 - 1$, ($r_0 \geq 2$), we obtain the uniform bound $\mathcal{O}(\delta_n^{2r+2})$ by uniform continuity of $\mathbb{L}^{(r+1, r+1)}(\cdot, \cdot)$ and results of Lemma 6.2. For $r = r_0, \dots, m_n$, we have $r^* = r_0$ so we apply the Hölderian regularity condition (6.9). Since $L_{i,k,r}(t) \leq r^r$, we arrive at $\sup_{t \in [0, T]} \mathbb{E}(X(t) - \tilde{X}_{r_0}(t))^2 = \mathcal{O}(\delta_n^{2(r_0 + \beta_0)})$ for $r = r_0$ while if $r = r_0 + 1, \dots, m_n$, $\sup_{t \in [0, T]} \mathbb{E}(X(t) - \tilde{X}_{\bar{r}}(t))^2 = \mathcal{O}(m_n^{2(m_n + r_0 + \beta_0)} \delta_n^{2(r_0 + \beta_0)})$. The logarithmic order of m_n yields the final result. In the case where $r_0 = 0$, above results hold true starting from

$$\begin{aligned} \mathbb{E}(X(t) - \tilde{X}_{\bar{r}}(t))^2 &= \sum_{i,j=0}^{\bar{r}} L_{i,k,\bar{r}}(t) L_{j,k,\bar{r}}(t) \left\{ \mathbb{L}(t, t) - \mathbb{L}(t, t_{k\bar{r}+j}) \right. \\ &\quad \left. - \mathbb{L}(t_{k\bar{r}+i}, t) + \mathbb{L}(t_{k\bar{r}+i}, t_{k\bar{r}+j}) \right\}. \end{aligned}$$

(b) For $e_\rho^2(\text{int}(\hat{r}_0))$, $\int_0^T (X(t) - \tilde{X}_{r_0+1}(t)) \rho(t) dt$ is again a Gaussian variable, so in a similar way as for approximation, we get the following bound for this term:

$$\begin{aligned} &\sqrt{3}(\mathbb{P}(\hat{r}_0 \neq r_0))^{\frac{1}{2}} \sum_{r=0}^{m_n} \left(\sup_{t \in [0, T]} \left(\mathbb{E}(X(t) - \tilde{X}_{r+1}(t))^2 \right)^{\frac{1}{2}} \right)^2 \left(\int_0^T \rho(t) dt \right)^2 \\ &+ \sum_{k=0}^{\lfloor \frac{n}{r_0+1} \rfloor - 1} \sum_{\ell=0}^{\lfloor \frac{n}{r_0+1} \rfloor - 1} \int_{\mathcal{I}_k} \int_{\mathcal{I}_\ell} \mathbb{E}(X(t) - \tilde{X}_{r_0+1}(t)) (X(s) - \tilde{X}_{r_0+1}(s)) \rho(t) \rho(s) ds dt. \end{aligned}$$

Study of the term $\mathbb{E}(X(t) - \tilde{X}_{r_0+1}(t)) (X(s) - \tilde{X}_{r_0+1}(s))$, $(s, t) \in \mathcal{I}_\ell \times \mathcal{I}_k$. Denoting $\bar{r} = r_0 + 1$ we get again from lemma 4.1 of Blanke and Vial (2008) that $\mathbb{E}(X(t) - \tilde{X}_{\bar{r}}(t)) (X(s) - \tilde{X}_{\bar{r}}(s))$ is equal to:

$$\begin{aligned} &\sum_{i,j=0}^{\bar{r}} L_{i,k,\bar{r}}(t) L_{j,\ell,\bar{r}}(s) \frac{((t_{k\bar{r}+i} - t_{k\bar{r}})(t_{\ell\bar{r}+j} - t_{\ell\bar{r}}))^{r_0}}{((r_0 - 1)!)^2} \iint_{[0,1]^2} dv dw ((1-v)(1-w))^{r_0-1} \\ &\times \left\{ \mathbb{L}^{(r_0, r_0)}(t_{k\bar{r}} + (t - t_{k\bar{r}})v, t_{\ell\bar{r}} + (t - t_{\ell\bar{r}})w) - \mathbb{L}^{(r_0, r_0)}(t_{k\bar{r}} + (t - t_{k\bar{r}})v, t_{\ell\bar{r}} + (t_{\ell\bar{r}+j} - t_{\ell\bar{r}})w) \right. \\ &\left. - \mathbb{L}^{(r_0, r_0)}(t_{k\bar{r}} + (t_{k\bar{r}+i} - t_{k\bar{r}})v, t_{\ell\bar{r}} + (t - t_{\ell\bar{r}})w) + \mathbb{L}^{(r_0, r_0)}(t_{k\bar{r}} + (t_{k\bar{r}+i} - t_{k\bar{r}})v, t_{\ell\bar{r}} + (t_{\ell\bar{r}+j} - t_{\ell\bar{r}})w) \right\}. \end{aligned}$$

For non-overlapping intervals \mathcal{I}_k and \mathcal{I}_ℓ , that is $|k - \ell| \geq 2$, we make use of Condition A2.2(2) four times, by adding and subtracting the necessary terms, noting that

$$\sum_{i,j=0}^{\bar{r}} L_{i,k,\bar{r}}(t) L_{j,\ell,\bar{r}}(s) (t_{k\bar{r}+i} - t_{k\bar{r}})^{r_1} (t_{\ell\bar{r}+j} - t_{\ell\bar{r}})^{r_2} = (t - t_{k\bar{r}})^{r_1} (s - t_{\ell\bar{r}})^{r_2}.$$

with either $r_i = \bar{r} - 1$ or $r_i = \bar{r}$ for $i = 1, 2$. By this way, we get

$$\begin{aligned} \sum_{\substack{k, \ell=0 \\ |k-\ell| \geq 2}}^{\lfloor \frac{n}{\bar{r}} \rfloor - 1} \int_{\mathcal{I}_k} \int_{\mathcal{I}_\ell} \mathbb{E}(X(t) - \tilde{X}_{\bar{r}}(t))(X(s) - \tilde{X}_{\bar{r}}(s)) \rho(t) \rho(s) \, ds dt \\ = \mathcal{O}\left(\delta_n^{2(r_0 + \beta_0 + 1)} \sum_{\substack{k, \ell=0 \\ |k-\ell| \geq 2}}^{\lfloor \frac{n}{\bar{r}} \rfloor - 1} |k - \ell - 1|^{-2(2 - \beta_0)}\right) \end{aligned}$$

which is a $\mathcal{O}\left(\delta_n^{2(r_0 + \beta_0 + 1)}\right)$. For overlapping intervals \mathcal{I}_k and \mathcal{I}_ℓ , that is in the case where $|k - \ell| \leq 1$, we make use of Cauchy-Schwarz inequality to obtain the same bound as above. Since the second part of $e_\rho^2(\text{int}(\hat{r}_0))$ is negligible, we obtain the result. \square

REFERENCES

- Adler, R. J. (1981). *The geometry of random fields*. Wiley, New-York.
- Adler, R. J. (1990). *An introduction to continuity, extrema, and related topics for general Gaussian processes*. Institute of Mathematical Statistics Lecture Notes—Monograph Series, 12. Hayward, CA: Institute of Mathematical Statistics.
- Benhenni, K. (1998). Approximating integrals of stochastic processes: extensions. *J. Appl. Probab.* 35(4), 843–855.
- Benhenni, K. and S. Cambanis (1992). Sampling designs for estimating integrals of stochastic processes. *Ann. Statist.* 20(1), 161–194.
- Berman, S. (1974). Sojourns and extremes of Gaussian processes. *Ann. Probab.* 2, 999–1026 (Corrections (1980), 8, 999 and (1984) 12, 281).
- Blanke, D. and C. Vial (2008). Assessing the number of mean-square derivatives of a Gaussian process. *Stochastic Process. Appl.* 118(10), 1852–1869.
- Blanke, D. and C. Vial (2011). Estimating the order of mean-square derivatives with quadratic variations. *Stat. Inference Stoch. Process.* 14(1), 85–99.
- Blanke, D. and C. Vial (2012). On estimation of regularity for Gaussian processes. *Preprint arXiv 1211.2763* (November), 34 pages. <http://arxiv.org/pdf/1211.2763v1>.
- Brown, P. J., T. Fearn, and M. Vannucci (2001). Bayesian wavelet regression on curves with application to a spectroscopic calibration problem. *J. Amer. Statist. Assoc.* 96(454), 398–408.
- Bucklew, J. A. (1985). A note on the prediction error for small time lags into the future. *IEEE Trans. Inform. Theory* 31(5), 677–679.
- Cambanis, S. (1985). Sampling designs for time series. In *Time series in the time domain*, Volume 5 of *Handbook of Statist.*, pp. 337–362. Amsterdam: North-Holland.
- Chan, G., P. Hall, and D. Poskitt (1995). Periodogram-based estimators of fractal properties. *Ann. Statist.* 23(5), 1684–1711.
- Constantine, A. G. and P. Hall (1994). Characterizing surface smoothness via estimation of effective fractal dimension. *J. Roy. Statist. Soc., Ser. B* 56(1), 97–113.
- Cressie, N. A. C. (1993). *Statistics for spatial data*. New-York: Wiley.
- Cuzick, J. (1977). A lower bound for the prediction error of stationary Gaussian processes. *Indiana Univ. Math. J.* 26(3), 577–584.
- Feuerverger, A., P. Hall, and A. Wood (1994). Estimation of fractal index and fractal dimension of a Gaussian process by counting the number of level crossing. *J. Time Ser. Anal.* 15(6), 587–606.

- Gilbert, R. O. (1987). *Statistical methods for environmental pollution monitoring*. New York: Van Nostrand-Reinhold.
- Gneiting, T., H. Ševčíková, and D. B. Percival (2012). Estimators of fractal dimension: assessing the roughness of time series and spatial data. *Statist. Sci.* 27(2), 247–277.
- Hall, P. and R. Roy (1994). On the relationship between fractal dimension and fractal index for stationary stochastic processes. *Ann. Appl. Probab.* 4(1), 241–253.
- Hanson, D. L. and F. T. Wright (1971). A bound on tail probabilities for quadratic forms in independent random variables. *Ann. Math. Statist.* 42(3), 1079–1083.
- Istas, J. (1992). Wavelet coefficients of a Gaussian process and applications. *Ann. Inst. H. Poincaré Probab. Statist.* 28(4), 537–556.
- Istas, J. and G. Lang (1997). Quadratic variations and estimation of the local hölder index of a Gaussian process. *Ann. Inst. H. Poincaré, Probab. Statist.* 33(4), 407–436.
- Istas, J. and C. Laredo (1997). Estimating functionals of a stochastic process. *Adv. in Appl. Probab.* 29(1), 249–270.
- Kent, J. T. and A. T. Wood (1997). Estimating the fractal dimension of a locally self-similar Gaussian process by using increments. *J. Roy. Statist. Soc., Ser. B* 59(3), 679–699.
- Laslett, G. M. (1994). Kriging and splines: an empirical comparison of their predictive performance in some applications. *J. Amer. Statist. Assoc.* 89(426), 391–409. With comments and a rejoinder by the author.
- Lindgren, G. (1979). Prediction of level crossings for normal processes containing deterministic components. *Adv. in Appl. Probab.* 11(1), 93–117.
- Müller-Gronbach, T. (1996). Optimal designs for approximating the path of a stochastic process. *J. Statist. Plann. Inference* 49(3), 371–385.
- Müller-Gronbach, T. and K. Ritter (1997). Uniform reconstruction of Gaussian processes. *Stochastic Process. Appl.* 69(1), 55–70.
- Müller-Gronbach, T. and K. Ritter (1998). Spatial adaption for predicting random functions. *Ann. Statist.* 26(6), 2264–2288.
- Plaskota, L., K. Ritter, and G. Wasilkowski (2004). Optimal designs for weighted approximation and integration of stochastic processes on $[0, \infty)$. *J. Complexity* 20(1), 108–131.
- R Core Team (2012). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.
- Ritter, K. (1996). Asymptotic optimality of regular sequence designs. *Ann. Statist.* 24(5), 2081–2096.
- Ritter, K. (2000). *Average-case analysis of numerical problems*. Lecture Notes in Mathematics, 1733. Springer.
- Sacks, J. and D. Ylvisaker (1968). Designs for regression problems with correlated errors; many parameters. *Ann. Math. Statist.* 39, 49–69.
- Sacks, J. and D. Ylvisaker (1970). Designs for regression problems with correlated errors. III. *Ann. Math. Statist.* 41, 2057–2074.
- Seleznev, O. (1996). Large deviations in the piecewise linear approximation of Gaussian processes with stationary increments. *Adv. in Appl. Probab.* 28(2), 481–499.
- Seleznev, O. (2000). Spline approximation of random processes and design problems. *J. Statist. Plann. Inference* 84(1-2), 249–262.
- Seleznev, O. and A. Buslaev (1998). Best approximation for classes of random processes. Technical Report 13, 14 p., Univ. Lund Research Report. <http://mech.math.msu.su/~seleznev/bestapp.ps>.

- Stein, M. L. (1995). Predicting integrals of stochastic processes. *Ann. Appl. Probab.* 5(1), 158–170.
- Stein, M. L. (1999). *Interpolation of spatial data*. Springer Series in Statistics. New York: Springer-Verlag. Some theory for Kriging.
- Taylor, C. C. and S. J. Taylor (1991). Estimating the dimension of a fractal. *J. Roy. Statist. Soc. Ser. B* 53(2), 353–364.
- Tsai, H. and K. S. Chan (2000). A note on the covariance structure of a continuous-time process. *Statist. Sinica* 10, 989–998.
- Wood, A. T. and G. Chan (1994). Simulation of stationary Gaussian processes in $[0, 1]^d$. *J. Comput. Graph. Statist.* 3(4), 409–432.

AVIGNON UNIVERSITÉ, LMA EA2151, 33 RUE LOUIS PASTEUR, F-84000 AVIGNON, FRANCE.
E-mail address: `delphine.blanke@univ-avignon.fr`

UNIVERSITÉ DE LYON, CNRS UMR 5208, POLYTECH LYON-UNIVERSITÉ DE LYON 1, INSTITUT CAMILLE JORDAN, 43 BLVD DU 11 NOVEMBRE 1918, 69622 VILLEURBANNE CEDEX, FRANCE.
E-mail address: `celine.vial@univ-lyon1.fr`